Predicting Reading and Mathematics From Neural Activity for Feedback Learning

Sabine Peters, Mara Van der Meulen, Kiki Zanolie, and Eveline A. Crone
Leiden University and Leiden Institute for Brain and Cognition, Leiden, the Netherlands

Although many studies use feedback learning paradigms to study the process of learning in laboratory settings, little is known about their relevance for real-world learning settings such as school. In a large developmental sample ($N = 228$, 8–25 years), we investigated whether performance and neural activity during a feedback learning task predicted reading and mathematics performance 2 years later. The results indicated that feedback learning performance predicted both reading and mathematics performance. Activity during feedback learning in left superior dorsolateral prefrontal cortex (DLPFC) predicted reading performance, whereas activity in presupplementary motor area/anterior cingulate cortex (pre-SMA/ACC) predicted mathematical performance. Moreover, left superior DLPFC and pre-SMA/ACC activity predicted unique variance in reading and mathematics ability over behavioral testing of feedback learning performance alone. These results provide valuable insights into the relationship between laboratory-based learning tasks and learning in school settings, and the value of neural assessments for prediction of school performance over behavioral testing alone.

Keywords: feedback learning, reading, mathematics, longitudinal, school performance

Learning from performance feedback is an important skill allowing us to rapidly adjust behavior based on changes in environmental demands (Holroyd & Coles, 2002). It is an adaptive form of learning allowing people to flexibly and creatively adapt to a changing environment. Feedback learning is often investigated in controlled laboratory settings to study the process of learning. However, it is currently unclear how feedback learning in these controlled experimental paradigms relate to real-world learning in settings, such as school. In this study, we investigated this question in a large developmental sample of participants between 8 and 27 years, focusing on both neural and behavioral indices of feedback learning as predictors for school performance 2 years later.

School performance can be measured in different ways. The most important school performance skills taught in schools across the world are reading and mathematics, of which reading is arguably the most important skill, given that many courses in school rely on children's ability to read proficiently. Also, many children who are poor readers in school keep having difficulties with reading later in life (O'Shaughnessy, Lane, Gresham, & Beebe-Frankenberger, 2003). Mathematical abilities are also important for many different subjects in school. More importantly, research has demonstrated that performance on mathematical tests predicts employability, productivity and salaries in adulthood (Geary, 2000; Rivera-Batiz, 1992).

One of the main reasons why laboratory-based feedback learning tasks and school performance may be related is because the capacity to learn from feedback is crucial in educational settings. In educational contexts, a large proportion of learning occurs based on the feedback from teachers, or feedback in other forms such as grades for tests or peer feedback. The degree to which children learn from feedback (which can be studied in an isolated way in laboratory settings) may predict how well they do on school tests as well, and is thus ultimately relevant for school performance. In the current feedback learning task participants were instructed to sort series of three pictures in a certain order. The sort order of the three pictures needed to be inferred based on positive and negative performance feedback. We distinguished between a "learning phase" and an "application phase" to study neural responses to feedback that is relevant for learning (early in the learning process), compared with feedback that is no longer informative because the participant already knows the correct location (later in the learning process). We tested whether the degree to which participants could successfully use feedback, was predictive of school learning measures. With regard to neural activity, we were specifically interested in neural responses to feedback during the learning phase compared with feedback during the application phase, as this may be an index of sensitivity to learning signals.

Aside from sensitivity to learning signals, another possible reason for a link between laboratory-based feedback learning tasks and school performance is that both feedback learning and reading and mathematics are linked to executive functions. Executive functions are defined as the ability to perform goal-directed actions in new situations and to overcome automatic thoughts and behaviors (Garon, Bryson, & Smith, 2008). Executive functions are thought to consist of three subprocesses, or basic executive functions: (a) working memory, (b) inhibition, and (c) switching (Huizinga, Dolan, & van der Molen, 2006; Miyake et al., 2000). It has been argued that complex cognitive tasks which rely on multiple

subprocesses of executive functions, such as the classic Wisconsin Card Sorting Task, are the most reliable correlates of cognitive challenges in daily life (Barcelo & Knight, 2002), possibly because these tasks are more similar to everyday challenges. Similar to the Wisconsin Card Sorting Test (WCST), the current feedback learning task relied on multiple aspects of executive functioning. For instance, participants needed working memory skills to keep relevant information online, they needed to plan for next trials, to form and test hypothesis on the correct sort order, they needed to switch hypotheses about the correct order after negative feedback, and so forth. Compared with the WCST, we aimed to study behavioral and neural reactions to learning from feedback per se, without the component of rule switches. Unexpected negative feedback after a string of correct answers, may result in very different brain activity compared with negative feedback during the search for a new correct rule (Barcelo & Knight, 2002). Thus, we focused on one aspect of the WCST which is rule learning, but not on the second aspect which is rule switching.

Several studies have provided evidence for a relationship between school performance and executive functioning. For instance, numerous studies have demonstrated a link between working memory, inhibition and switching on the one hand, and reading and mathematics performance on the other (Blair & Razza, 2007; Bull & Scerif, 2001; Raghubar, Barnes, & Hecht, 2010; Van der Sluis, De Jong, & Van der Leij, 2004). The link between executive functioning and school performance is not surprising, given that to develop reading and mathematics understanding, children probably need additional cognitive skills. For example, children have to be able to understand grammatical and numerical structure, keep track of the sentences read or mathematical steps taken before, and integrate information from long-term memory with current information to form a coherent view (Cain, Oakhill, & Bryant, 2004; Landi, Frost, Mencl, Sandak, & Pugh, 2013), which are all processes intimately related to executive functioning. This led us to hypothesize that feedback learning in controlled laboratory settings is a valid predictor of real-world learning performance in schools.

Recently, an increasing body of research has directed attention to predicting school performance from brain measures. A possible advantage of collecting neural measures in addition to behavioral measures is the hypothesis that brain measures can provide unique predictive information over behavioral measures alone. A prior study already demonstrated that neural activity during a working memory task predicted unique variation in mathematical performance 2 years later (Dumontheil & Klingberg, 2012). Similarly, reading encoding ability was better predicted by a combination of neural and behavioral measures rather than behavioral testing alone (Hoeft et al., 2007). The main neural areas involved during feedback processing are the dorsolateral prefrontal cortex (DLPFC), superior parietal cortex (SPC), and presupplementary motor area/anterior cingulate cortex (pre-SMA/ACC; Peters, Braams, Raijmakers, Koolschijn, & Crone, 2014; Zanolie, Van Leijenhorst, Rombouts, & Crone, 2008). Meta-analyses of fMRI-activity during reading and mathematics also show recruitment of these areas (Arsalidou & Taylor, 2011; Ferstl, Neumann, Bogler, & Von Cramon, 2008; Houdé, Rossi, Lubin, & Joliot, 2010) and pre-SMA/ACC (Ferstl et al., 2008; Houdé et al., 2010) among other areas (mostly lateralized to the left hemisphere). Interestingly, meta-analyses on mathematics-related neural activity also showed

involvement of the DLPFC (Arsalidou & Taylor, 2011; Houdé et al., 2010), parietal cortex and pre-SMA/ACC (Arsalidou & Taylor, 2011). Possibly, the same neural regions that respond to feedback signals are also associated with reading and mathematics.

In this study, we investigated the link between learning in a controlled laboratory setting, and reading and mathematical ability as indices for real-world learning. We focused on fluency at reading single words, because this is one of the most crucial aspects of reading determining reading ability at a later stage (Jenkins, Fuchs, Van Den Broek, Espin, & Deno, 2003; Juel, 1988). To assess mathematics proficiency, we used a standardized arithmetic test that is part of the Wechsler Adult Intelligence Scale and the Wechsler Intelligence Scale for Children. This task measures numerical reasoning and mathematical problem solving and relies on the use of mathematical facts, procedures, and concepts, all of which have been related to executive functioning skills (Cragg & Gilmore, 2014). In addition, we investigated whether individual differences in working memory could explain a possible link between feedback learning and reading and mathematics performance. For instance, Huizinga, Dolan, and van der Molen (2006) found in a sample of 7- to 21-year-olds that from the factors working memory, inhibition, and switching, only working memory predicted WCST performance, a task that also relies on learning from feedback. We hypothesized that feedback learning would predict reading and mathematics performance 2 years later, and that neural measures would provide additional information over behavioral testing (feedback learning performance, working memory) alone.

Method

Participants

The initial sample consisted of 299 participants (data also published in Peters, Braams, et al., 2014; Peters, Koolschijn, Crone, Van Duijvenvoorde, & Raijmakers, 2014), for whom data was collected on two time points (T1 and T2) which were approximately 2 years apart ($M = 1.99$, $SD = 0.10$, range: 1.66–2.47 years). The included sample with complete data at T1 for feedback learning and fMRI data consisted of 268 participants. At T1 participants were excluded from analyses for a variety of reasons, such as reported history of neurological or psychiatric disorders or use of psychotropic medication, movement in the MRI scanner exceeding 3.0 mm ($N = 19$), technical issues ($N = 3$), or because they were outliers at the lower end (more than three times the interquartile range) on feedback learning performance ($N = 3$).

Reading and mathematics data was only obtained at T2. At T2, there was complete data on reading and math performance for 228 participants (119 females) who were also included at T1 (aged 8.01–24.55 years at T1 ($M = 14.35$, $SD = 3.57$) and aged 9.92–26.62 at T2 ($M = 16.34$, $SD = 3.58$)). All analyses were performed on these 228 participants (see Table 1 for an overview of these participants per age and per sex). IQ scores at T1 were estimated using two subtests (similarities and block design) of the WISC–III (participants 8- to 15-years-old) or WAIS–III (participants 16- to 25-years-old). Estimated IQ scores ranged from 85 to 143 ($M = 110.78$, $SD = 9.80$). The study was approved by the Institutional Review Board at the University Medical Center and all participants older than 12 years of age (and participants' parents

Table 1
Number of Participants (Final Included Sample) Per Age and Per Sex

Age	Sex		
	Male	Female	Total
8 years	6	4	10
9 years	10	5	15
10 years	8	8	16
11 years	8	11	19
12 years	16	9	25
13 years	16	15	31
14 years	10	14	24
15 years	10	11	21
16 years	10	8	18
17 years	11	11	22
18–25 years	14	13	27
Total	119	109	228

for children under 18 years of age) signed an informed consent form. Adults received payment (€60) for participation and children and their parents received brain-related presents and a payment for travel reimbursement (€30 for children 12–17 years, €25 for children 8–11 years).

Materials

Reading fluency. Technical reading skills were measured at T2 with a reading fluency task. We used one of the tests in the Dutch "Three-Minute-Test" (Krom, Jongen, Verhelst, Kamphuis, & Kleintjes, 2010). In this task, participants received a list of words and were instructed to read aloud as many words as possible in 1 min. The total score is defined as the number of correct words read minus the number of incorrect words. The Three-Minute-Test has good validity and reliability (Cronbach's alpha, dependent on age group >0.92; Krom et al., 2010).

Mathematics. Mathematical ability was measured at T2 with the subscale "arithmetic" of the Wechsler Intelligence Scales (WISC–III for participants under 16 years of age, WAIS-III for participants of 16 years and older). A set of arithmetical problems of increasing difficulty was administered verbally. All arithmetic problems had a time limit of 30 s to 75 s, depending on the difficulty of the problem. If the participants failed to correctly answer three consecutive problems the test was aborted. Both the WISC and the WAIS resulted in raw scores that were converted to norm scores relative to same-aged peers. We used norm scores in further analyses (see also Barnea-Goraly et al., 2005; Li, Hu, Wang, Weng, & Chen, 2013) to ensure comparability between the different ages (reflected in WISC and WAIS scores). In addition, we performed our main analyses with the mathematics subtest with raw scores for the WISC and WAIS group separately.

Working memory. We measured working memory performance at T1 to assess whether feedback learning and reading and mathematics performance were explained by individual differences in working memory. Working memory capacity was measured with the Mental Counters task (Huizinga et al., 2006), in which participants need to keep numerical information active. For this task, two independent counters were presented on a computer screen. The counters were horizontal bars for which the values

changed depending on the position of a square. If a square was presented above a counter the participant was instructed to add 1 to the current value, if a square was presented below the counter the participant was instructed to subtract 1 from the current value of the counter. The squares appeared randomly above or below one of the two counters. Participants were explicitly instructed to use a verbal counting strategy, by instructing them to keep track of both counters by mentally counting (e.g., 0–0, 1–0, 1–1, 1–2, etc.) and to press a button as soon as one of the counters reached a given criterion value (e.g., when one of the counters reached the value 3). The squares were randomly presented in series (the number of trials before criterion was reached) of five or seven trials with intertrial intervals of 1,000 to 1,300 ms, with a total of 16 trials. The proportion of correct trials was used as a measure of performance.

Feedback learning task. Participants performed a feedback learning task in the MRI scanner (Peters, Braams, et al., 2014; Peters, Koolschijn, et al., 2014). On every trial, three empty boxes were presented in the top half of the screen in the stimulus and feedback display. During presentation of the stimulus display one of three different stimuli was presented in the center of the bottom half of the screen (see Figure 1). Participants were instructed that each stimulus belonged in one of three boxes for an entire sequence and they had to find the correct location for all three stimuli by using performance feedback. Each trial started with a 500 ms fixation cross, presented in the center of the screen. After fixation the stimulus display was presented for 2,500 ms, during which participants were required to sort the stimulus in one of three squares. Participants responded by pressing one of three buttons strapped to their right leg. If participants failed to respond within 2,500 ms "Too Late" was presented in the center of the screen, after which the sequence continued. After the response, performance feedback was presented for 1,000 ms. When a participant sorted a stimulus in the correct square a "plus" sign (positive feedback) was shown, when a participant sorted a stimulus in the incorrect square a "minus" sign (negative feedback) was shown. Intertrial interval (blank screen) was jittered to optimize the timing for fMRI based on OptSeq (Dale, 1999) with intervals between 0 s and 6 s. A sequence was aborted when the participant sorted each stimulus twice in the correct location, or after 12 trials in total. When a sequence ended a new sequence with three new unique stimuli was presented. There were 15 sequences in total, resulting in a maximum of 180 trials. The mean number of trials was 138.80
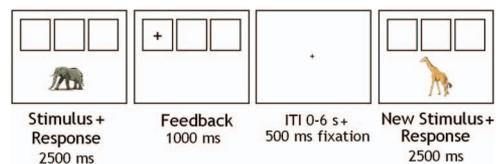

Figure 1. Display of task sequence for the feedback learning task. A trial started with a 2,500-ms stimulus display during which the participant responded by sorting the stimulus in one of the three boxes. In this example, the participant (correctly) chose the left box. Next, feedback was presented for 1,000 ms by either a "+" for correct feedback or a "−" for incorrect feedback. After an intertrial interval (varying from 0–6 s) and a 500-ms fixation cross, the next stimulus was presented. See the online article for the color version of this figure.

(SD = 9.25, range = 117–165 trials). Stimuli were presented in a pseudorandom order, with a maximum of two identical stimuli in a row. During the MRI session the task was divided into two runs of eight and seven sequences, respectively. Before the MRI session, all participants practiced three sequences. During the practice session, experimenters were observing the participants' responses to check whether the participant had understood the task. Almost all participants understood the instructions. If not, further instructions were provided to make sure all participants were able to perform the task.

To calculate a performance measure for feedback learning we calculated the percentage of trials in the learning phase where feedback was successfully used on the next trial. That is, if a participant received positive feedback for a stimulus, and opted for the same response the next time that stimulus appeared, this would mean that the positive feedback was successfully applied in a next trial. For negative feedback, feedback was successfully applied when the participant did not choose the same option in a next trial. To calculate the performance measure, we divided the number of trials during the learning phase which were successfully applied in the next trial, by the total number of trials during the learning phase.

fMRI Data Acquisition

MRI scans were obtained with a Philips 3.0 Tesla MRI scanner. Functional scans for the feedback learning tasks were acquired during two runs with T2*-weighted echo-planar imaging (EPI). The first two volumes were discarded to allow for equilibration of T1 saturation effects. The following settings were used: TR = 2.2 s, TE = 30 ms, sequential acquisition, 38 slices, slice thickness = 2.75 mm, Field of View (FOV) = 220 mm × 220 mm × 114.68 mm. For the structural scan, a high-resolution 3D T1-FFE was obtained after the experimental tasks (TR = 9.76 ms, TE = 4.59 ms, 140 slices, voxel size = 0.875 mm, FOV = 224 mm × 177 mm × 168 mm). The experimental task was projected on a screen, which was visible to participants through a mirror. Total scan duration for the task was on average 11.57 min (range 9.75–13.75 min). Participants were accustomed to the MRI environment and sounds with a mock scanner before the actual MRI scan.

fMRI Data Analysis

We used SPM8 (Wellcome Department of Cognitive Neurology, London, United Kingdom) to analyze fMRI. The following preprocessing steps were used: correction for slice timing acquisition and rigid body motion, spatial normalization to T1 templates (MNI305 stereotaxic space; Cocosco, Kollokian, Kwan, & Evans, 1997) using a 12-parameter affine transform together with a nonlinear transformation involving cosine basis functions and resampling of the volumes to 3 mm voxels. Functional scans were smoothed with an 8-mm FWHM isotropic Gaussian kernel. For further fMRI analyses, we used a contrast that reveals brain areas with sensitivity to informative feedback for learning (Eliassen et al., 2012; van den Bos, Güroğlu, van den Bulk, Rombouts, & Crone, 2009), that is, areas responding more to feedback providing new information (i.e., more informative) compared with feedback providing known information. To compare neural activity for

"informative" and "uninformative" feedback, we distinguished between a learning phase and an application phase for each stimulus. For the learning phase, we included trials where participants had not correctly sorted this particular stimulus yet, and were thus still using feedback to determine the correct location. Only trials for which feedback was used appropriately on the next trial for that stimulus were included. Thus, feedback was categorized as learning, when positive feedback resulted in choosing the same location on a next trial and when negative feedback resulted in sorting in a different location. These trials during the learning phase were compared with the application phase: trials in which a stimulus was sorted correctly on a preceding trial, and continued to be sorted correctly. All further analyses were based on a comparison between the learning phase and the application phase, that is, the contrast Learning > Application. In order to calculate this contrast for all participants, we first modeled the fMRI time series with events corresponding to the events "positive learning," "negative learning," and "application," time-locked with 0-duration to the moment of feedback, which were convolved with a canonical hemodynamic response function. Other trials (e.g., trials during the learning phase that did not result in learning or trials where participants responded too late) were modeled as events of no interest. The events were used in a general linear model; along with a set of cosine functions which high-pass filtered the data. The least-squares parameter estimates of height of the best-fitting canonical HRF for each condition were used for the calculation of the contrast Learning (Positive Learning + Negative Learning) > Application for each subject. The resulting contrast images were submitted to higher-level analyses.

fMRI Region-of-Interest Analysis

In order to examine neural effects of feedback learning and its relation to reading and mathematics performance, region-of-interest (ROI) analyses were performed with the Marsbar toolbox in SPM8 (Brett, Anton, Valabregue, & Poline, 2002). The contrast used to generate functional ROIs was Learning > Application (FWE corrected, $p < .05$, >10 contiguous voxels). The resulting ROIs spanned several brain regions. Therefore, the ROIs were subdivided by masking the functional ROI with the following anatomical Marsbar ROIs (based on automated anatomical labeling [AAL]), left and right DLPFC (middle frontal gyrus in AAL), pre-SMA/ACC (supplementary motor area in AAL; left and right combined), and left and right SPC (superior parietal lobule in AAL). These ROIs were selected based on earlier studies demonstrating that these areas show developmental changes for feedback learning (Crone, Zanolie, Van Leijenhorst, Westenberg, & Rombouts, 2008; Peters, Braams, et al., 2014; van Duijvenvoorde, Zanolie, Rombouts, Raijmakers, & Crone, 2008) and were also used in a prior study with the same experimental task (Peters, Braams, et al., 2014). The DLPFC ROIs, even after masking, were still very large (right: 28,488 mm; left: 282,40 mm), therefore, we created 6-mm radius spheres based on four local maxima within the DLPFC regions (two per hemisphere). These areas are referred to as "superior DLPFC (sup-DLPFC)" and "mid-DLPFC." Centre-of-mass MNI (x, y, z) coordinates for the ROIs were: pre-SMA/ACC: 0, 9, 58; right sup-DLPFC: 21, 9, 57; left sup-DLPFC: −24, 3, 57; right mid-DLPFC: 42, 18, 39;

left mid-DLPFC: −42, 24, 39; right SPC: 27, −62, 55; left SPC: −23, −64, 50 (see Figure 2).

Results

Data Checks

We performed several data quality checks by investigating relationships between the main variables of interest (neural activity and behavioral performance for feedback learning, and reading and mathematics), and age, IQ, working memory, and sex (see Table 2 for an overview of the values for age, IQ, working memory, feedback learning, reading, and mathematics). There were no sex differences in any of the measures except for working memory which was higher for males (see Table 2). Because there were no sex differences for the key variables reading, mathematics, and feedback learning, we did not investigate sex effects further. We also investigated correlations between the different variables. There was an age-corrected correlation between reading and mathematics scores, $r = .20$, $p = .003$. Working memory at T1 correlated positively (corrected for age) with feedback learning performance at T1, $r = .33$, $p < .001$; reading fluency at T2, $r = .15$, $p = .026$; and mathematics at T2, $r = .25$, $p < .001$ but not with neural activity at T1. IQ estimates at T1 correlated with mathematics norm scores at T2 ($r = .32$, $p < .001$, age-corrected) but not with the other measures (reading fluency, feedback learning, and neural activity).

With regard to age effects, we found that age at T1 correlated positively with reading fluency, $r = .31$, $p < .001$; working memory, $r = .34$, $p < .001$; and feedback learning performance, $r = .47$, $p < .001$. Age was also positively related to neural activity for the difference score Learning > Application in all seven ROIs. Therefore, we corrected for age in further analyses. Even though mathematics scores were norm scores, that is, scores relative to same-aged peers, there was still a small but significant correlation with age, $r = .16$, $p = .018$. We therefore also corrected for age in all further analyses with mathematics scores.
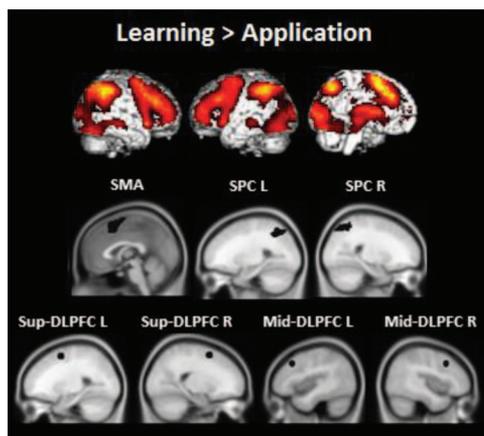

Figure 2. Wholebrain results for the contrast Learning > Application (FWE-corrected at $p < .05$, > 1 0 contiguous voxels) and the regions-of-interest based on this contrast. See the online article for the color version of this figure.

Figure 3 shows the relations with age separated in categories for illustrative purposes.

For the outcome variables (reading fluency and mathematics scores) we furthermore tested whether the relationship with age was best described by a linear function of age at T1, a quadratic function of age, or a cubic function of age (for a similar approach, see Braams, van Duijvenvoorde, Peper, & Crone, 2015; Somerville et al., 2013). We used a hierarchical regression approach with reading/mathematics as dependent variable, and added age[1] as first step, age[2] as second step, and age[3] as third step, in order to test whether polynomials of age explained the relation better compared with a simpler model. For mathematics, the relationship with age was best described by a linear age pattern. However, reading fluency showed a combined linear and quadratic age pattern, leveling off toward adulthood. In all further analyses with reading fluency as a dependent variable, we therefore added not only age[1] but also age[2] as a control variable.

Predicting Reading and Mathematics Performance at T2 From T1 Feedback Learning

We first investigated whether reading and mathematics performance at T2 could be predicted from behavioral performance on the feedback learning task at T1. A hierarchical regression with age at T1 entered as a first step and feedback learning performance at T1 as a second step, showed that in addition to age, feedback learning performance significantly predicted reading fluency and mathematics performance 2 years later (positive relation; see Table 3).

Predicting Reading and Mathematics Performance at T2 From T1 Neural Activity During Feedback Learning

Next, we assessed whether brain activity during feedback learning in seven ROIs at T1 predicted reading and mathematics performance at T2. We performed hierarchical regressions with age at T1 as first step and neural activity in one of the seven ROIs as second step. These analyses showed that in addition to age and age[2], reading fluency was predicted by left sup-DLPFC activity (see Table 4). For mathematics performance at T2, activity in pre-SMA/ACC was a significant predictor above age (see Table 5). For a visual representation of the relationship between right sup-DLPFC activity and mathematics performance, and left sup-DLPFC and reading fluency (see Figure 4).

We also tested whether neural activity for feedback learning explained additional variance in reading and mathematics above age and behavioral performance for feedback learning. We analyzed this with hierarchical regressions with age (and age[2] for reading fluency) at T1 as first step, feedback learning performance at T1 as second step, and neural activity (per ROI) as third step. Neural activity explained additional variance above behavioral measures for both reading fluency (left sup-DLPFC remained significant; $\beta = .15$, $p = .026$) and mathematics (pre-SMA/ACC remained significant; $\beta = .16$, $p = .022$). This indicates that neural activity in left sup-DLPFC and pre-SMA/ACC explained unique variance in reading and mathematics over and beyond age and behavioral feedback learning performance.

Table 2

Descriptive Values for Age, IQ, Working Memory, Feedback Learning, Reading and Mathematics Scores for Male and Female Participants Separately

	Female				Male				
	Mean	SD	Min	Max	Mean	SD	Min	Max	p sex
Age T1	14.10	3.39	8.01	22.79	14.63	3.75	8.01	24.55	.27
Age T2	16.10	3.40	10.02	24.83	16.60	3.77	9.92	26.62	.30
IQ T1	109.83	10.09	85.00	143.00	111.81	9.40	93.00	138.00	.13
Working Memory T1	.79	.17	.13	1.00	.86	.12	.38	1.00	$p < .001$
Feedback Learning T1	93.62	5.36	71.29	100.00	93.78	4.40	81.11	100.00	.81
Reading fluency T2	98.02	14.51	64.00	120.00	97.72	15.46	58.00	120.00	.88
Mathematics T2	11.75	2.88	6.00	19.00	12.44	2.69	4.00	18.00	.06

Note. In the right-most column, we indicated the *p*-value for sex differences.

Cross-Validation

To further confirm these models, we used leave-one-out cross-validation using the cv.glm function in R package boot (Canty & Ripley, 2016). This method leaves out one participant at every turn and predicts the dependent variable using all remaining data (*N*-1). These analyses indicated that the main results were confirmed by cross-validation. That is, for reading fluency a model including age, age^2 and feedback learning performance resulted in a lower prediction error (192.34) compared with a model including only age and age^2 (194.35). For mathematics norm scores, prediction error for a model including age and feedback learning (7.35) was lower compared with a model including only age (7.77). Also the neuroimaging results were confirmed using cross-validation. A model to predict reading fluency with age, age^2, and left superior DLPFC activity resulted in a lower prediction error (191.05)

compared with a model with only age and age^2 (194.35). This was also the case when we tested whether neural activity explained additional variance over feedback learning performance: a model with age, age^2, feedback learning performance, and left superior DLPFC activity resulted in a lower prediction error (189.76) compared with a model with age, age^2, and feedback learning performance (192.34). The model predicting mathematics from age and SMA activity resulted in a lower prediction error (7.59) than a model with only age as predictor (7.77). The cross-validation approach confirmed that SMA activity also explained additional variance above feedback learning performance, that is, a model with age, feedback learning performance, and SMA activity resulted in a lower prediction error (7.25) compared with a model with age and feedback learning performance (7.35).
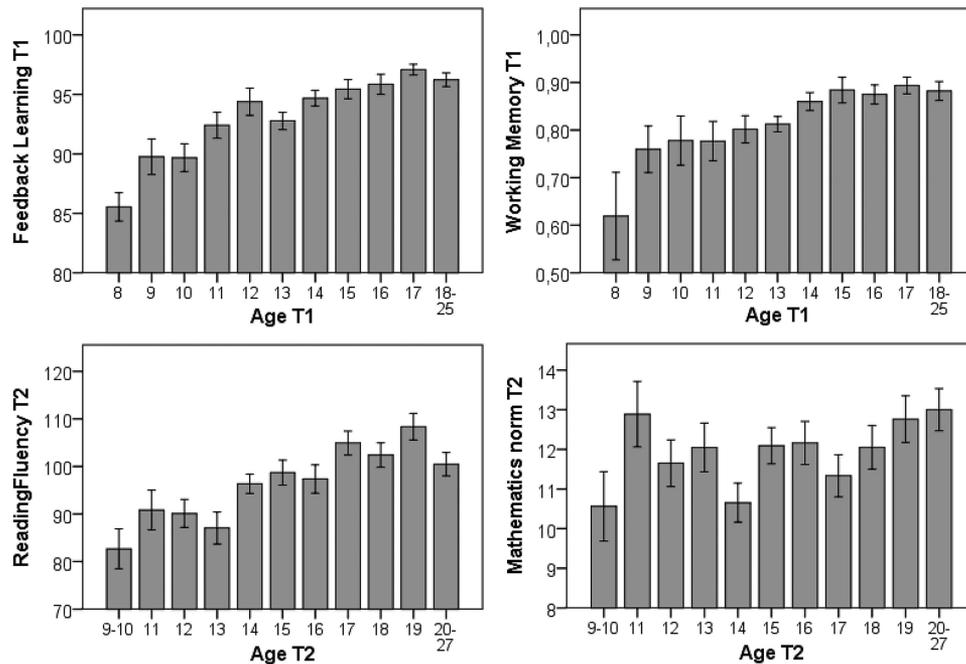

Figure 3. Display of age effects for feedback learning, working memory, reading, and mathematics. Note that for T2 one participant was 9.92-years-old, therefore the youngest age group at T2 was 9 and 10 years combined.

Table 3
Hierarchical Linear Regression Models With Age and Feedback Learning Performance as Significant Predictors for Reading and Mathematics Performance

Steps	Predictor	B	$SE\ B$	β	p	F	R^2
			Dependent: Reading fluency T2				
1	Overall model					19.334***	.15
	Age T1	7.68	1.72	1.83	<.001***		
	Age T1^2	−.21	.055	−1.55	<.001***		
2	Overall model					14.510***	.16
	Age T1	6.17	1.85	1.47**	.001**		
	Age T1^2	−.17	1.85	−1.25	.004**		
	Feedback learning T1	.46	.22	.15	.039*		
			Dependent: Mathematics T2				
1	Overall model					5.47*	.02
	Age T1	.12	.05	.15	.020*		
2	Overall model					10.53***	.09
	Age T1	.02	.06	.02	.760		
	Feedback learning T1	.16	.04	.28	<.001***		

* $p < .05$.   ** $p < .01$.   *** $p < .001$.

Specificity of the Effects for Reading Versus Mathematics

The correlation between reading and mathematics performance (while controlling for age) was significant, $r = .19$, $p = .003$. We therefore performed follow-up analyses to test whether the prediction of reading and mathematics from feedback learning was specific to either reading or mathematics, or whether the predictive effect may be related to a more general effect on school performance measures. The follow-up regression analyses showed that feedback learning performance no longer predicted reading fluency ($\beta = .08$, $p = .273$) over age and age^2 when adding mathematics performance to the model, indicating a general effect on school performance rather than a specific effect for reading. In contrast, feedback learning performance still predicted mathematics performance ($\beta = .13$, $p = .047$) in addition to age and reading fluency, suggesting a specific effect for mathematics.

Next, the same follow-up analyses were performed with neural activity as predictor for reading and mathematics. In the previous paragraph, we showed that left sup-DLPFC predicted reading fluency over age and age^2. Follow-up analyses indicated that this effect remained significant ($\beta = .14$, $p = .031$) when adding mathematics performance to the model, suggesting a specific re-

lation between sup-DLPFC activity and reading fluency. Also, pre-SMA/ACC ($\beta = .18$, $p = .011$) still predicted mathematics over age when adding reading fluency to the model, suggesting a specific effect on school performance measures.

Specificity of the Effects for Different Brain Regions

To test whether reading fluency and mathematics performance were specifically predicted by these two neural regions, or whether there was also a general effect of activation in the feedback learning network, we calculated mean activation across all seven ROIs. Next, we performed regression analyses with reading fluency and mathematics as dependent variables, age (and age^2 for reading) as first step, mean activity in all ROIs as second step, and left sup-DLPFC activity/pre-SMA/ACC activity as third step. The results showed that reading fluency was specifically predicted by left sup-DLPFC activity ($\beta = .24$, $p = .007$), and the mean activity across all ROIs was not a significant predictor ($\beta = −.13$, $p = .167$). For mathematics, mean activity for all ROIs was not a significant predictor either ($\beta = −.03$, $p = .821$), but activity in pre-SMA/ACC only remained marginally significant when adding mean activity in all ROIs to the model ($\beta = .21$, $p = .072$).

Table 4
Hierarchical Linear Regression Models for Neural Activity in Left Sup-DLPFC as Significant Predictor Above Age for Reading Fluency

Steps	Predictor	B	$SE\ B$	β	P	F	R^2
1	Overall model					19.33***	.15
	Age T1	7.68	1.72	1.83	<.001***		
	Age T1^2	−.21	.06	−1.55	<.001***		
2	Overall model					15.098***	.17
	Age T1	7.08	1.72	1.71	<.001***		
	Age T1^2	−.20	.06	−1.46	<.001***		
	Sup-DLPFC L	1.12	.47	.16	.017*		

*** $p < .001$.

Table 5
Hierarchical Linear Regression Models for Neural Activity in Pre-SMA/ACC as Significant Predictor Above Age for Mathematics Performance

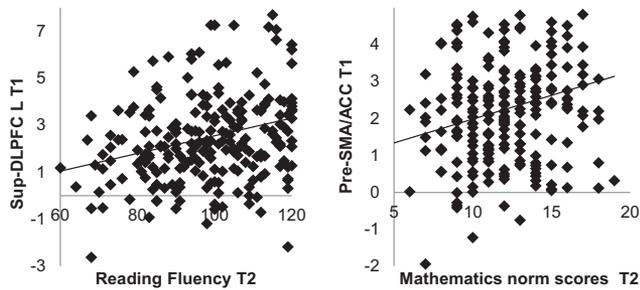| Predictor	B	$SE\ B$	β	p	F	R^2
		Dependent: Mathematics				
1. Overall model					5.47*	.02
Age T1	.12	.05	.15	.020*		
2. Overall model					6.31**	.05
Age T1	.08	.05	.10	.159		
Pre-SMA/ACC	.54	.21	.18	.009**		

* $p < .05$.   ** $p < .01$.

Figure 4. Scatterplot of the significant relationships between reading and mathematics performance at T2 and neural activity at T1 for the contrast Learning > Application.

Adding Working Memory and IQ as Control Variables

To assess whether the relationship between feedback learning and reading and mathematics performance could be explained by individual differences in working memory, we tested whether the above effects remained significant when analyzing a hierarchical regression with age (and age^2 for reading fluency) as a first step, working memory and IQ at T1 as a second step, and feedback learning performance or neural activity as a third step. Reading fluency was still predicted by left sup-DLPFC ($\beta = .15$, $p = .023$), over age ($\beta = 1.51$, $p = .001$), age^2 ($\beta = -1.31$, $p = .003$), IQ ($\beta = .029$, $p = .652$), and working memory ($\beta = .094$, $p = .173$). However, despite the fact that neither IQ ($\beta = .15$, $p = .877$) nor working memory ($\beta = .97$, $p = .335$) was a significant predictor of reading fluency, the prediction of reading fluency from feedback learning performance was no longer significant ($\beta = .13$, $p = .093$), which is possibly related to the increased number of predictors in the analysis.

For mathematics, feedback learning performance remained a significant predictor ($\beta = .18$, $p = .015$) over age ($\beta = .16$, $p = .018$), working memory ($\beta = .22$, $p = .001$), and IQ ($\beta = .30$, $p < .001$). Pre-SMA/ACC ($\beta = .16$, $p = .016$) was also still a significant predictor over age ($\beta = .051$, $p = .456$), IQ ($\beta = .29$, $p < .001$) and working memory ($\beta = .21$, $p = .002$). Together, these results indicate that feedback learning performance and neural activity explained unique variance in reading and mathematics that was not explained by working memory or IQ.

Mathematics Raw Scores

All prior analyses used mathematics norm scores. To investigate whether results were also present when using raw scores, we also performed the analyses with feedback learning performance and neural activity as predictors for raw mathematics scores. Because the younger age group (10–15, $N = 116$) performed the mathematics test from the WISC–III and the older group (16–27, $N = 112$) the WAIS-III, these age groups were analyzed separately. The results showed that effects were only present in the younger adolescents but not the in the older adolescent/adult group. That is, for the youngest group, mathematics performance was predicted above age by feedback learning performance ($\beta = .14$, $p = .027$) and by pre-SMA/ACC activity ($\beta = .23$, $p = .030$). None of the

effects were significant for the participants who were 16 years and older, suggesting a specific effect for the younger age range.

Discussion

In this study we investigated whether performance and neural activity during a feedback learning paradigm, used to study learning processes in a controlled laboratory setting, could predict indices of real-world learning performance in school 2 years later (reading and mathematics performance). The results of this study showed that (a) feedback learning performance predicted both reading and mathematics performance 2 years later; (b) neural activity during feedback learning in left sup-DLPFC predicted reading fluency, and neural activity in pre-SMA/ACC predicted mathematics performance 2 years later; (c) these neural regions predicted unique variation in school performance over behavioral testing alone; and (d) relations between feedback learning performance and neural activity with school performance remained significant when controlling for individual differences in working memory capacity and IQ. These results are discussed further in the following paragraphs.

Relation Between Feedback Learning Performance and School Performance

For both reading and mathematics, our results indicated that performance could be predicted by feedback learning performance 2 years earlier. To date, no prior research has investigated the relation between laboratory-based feedback learning measures and indices of real-world learning in school settings. The results confirmed our hypothesis that laboratory based learning measures can be powerful predictors of school outcomes. A possible reason for the relation between feedback learning and school performance measures is that the ability to learn from feedback is very important in school settings. Learning in schools relies in large part on learning from performance feedback from teachers or test scores. Children may differ in their ability to learn from feedback, and this may ultimately influence school performance.

It is also possible that the relation between feedback learning and school performance is explained by underlying individual differences in executive functions. It is well conceptualized that both feedback learning and school performance are related to executive functions (Diamond, 2013). Consistent with this, we found a positive correlation between working memory performance and feedback learning, as well as between working memory and reading and mathematics performance. However, even when adding working memory as a predictor to the model, feedback learning performance still predicted unique variance for both reading and mathematics, suggesting that working memory may explain a part of, but not all variance. Note that in our study, we only included a measure of working memory (but no measure of other executive functions). Another way to have a better understanding of the mechanisms through which feedback learning predicts school outcomes is by understanding the neural activity related to feedback learning, given that these are strongly linked to learning performance (Peters, Braams, et al., 2014).

Relation Between Neural Activity for Feedback Learning and School Performance

An important question tested in this study was whether neural activity could predict reading and mathematics performance 2 years later, and whether neural activity could provide additional information over behavioral testing alone. This was based on prior studies showing that neural measures can predict reading (Hoeft et al., 2007; Maurer et al., 2009) and mathematics performance (Dumontheil & Klingberg, 2012). Consistent with these studies, we found evidence for a relation between neural activity for feedback learning and reading and mathematics ability. First, we found that left sup-DLPFC activity predicted reading ability. These findings fit with earlier research showing that a mostly left-lateralized network including DLPFC is involved during reading tasks (Ferstl et al., 2008). Second pre-SMA/ACC predicted mathematics ability 2 years later. This fits with meta-analyses showing involvement of pre-SMA/ACC during arithmetical tasks (Arsalidou & Taylor, 2011; Houdé et al., 2010). Notably, for all areas we found a positive relation, indicating that increased activity predicts better performance on reading or mathematics tests. With the current design, it is not possible to determine whether higher activity might indicate better functioning or perhaps earlier maturation of these regions. Future research could build on this study by analyzing longitudinal fMRI measures and data on structural brain development.

In addition, we performed analyses to assess whether neural measures provided unique information that cannot be captured by behavioral testing alone. Both regions remained significant predictors when controlling for behavioral performance during feedback learning. This indicated that assessing feedback learning ability is useful for predicting reading and mathematics, but adding neural measures in addition to behavioral assessment further enhanced predictive ability. The finding that neural activity measures have added value over behavioral testing alone fits with earlier studies for the prediction of reading (Hoeft et al., 2007) and mathematics (Dumontheil & Klingberg, 2012)..

Prior research suggested that working memory is an important component of both feedback learning (Miyake et al., 2000) and reading and mathematics (Alloway & Alloway, 2010); therefore, it was possible that working memory is the underlying factor explaining these relations. When we controlled for working memory and IQ, there was still a significant prediction of reading fluency from feedback learning performance and activity in left sup-DLPFC, and for prediction of mathematics from feedback learning performance and activity in pre-SMA/ACC. This indicates that although working memory may play a role in the relation between feedback learning and reading and mathematics, there is still unique variation in reading and mathematics that is explained by neural activity during feedback learning. Other aspects of feedback learning performance that might be relevant for learning in school settings, are for instance the capacity to monitor one's actions and keep track of performance feedback, ignoring irrelevant aspects of the task, perceived competence and motivation (Fortier, Vallerand, & Guay, 1995; St. Clair-Thompson & Gathercole, 2006). Future research is needed to examine this in more detail.

Limitations and Future Directions

There are several limitations to this study. First, school performance can be measured in many ways. In this study, we measured only two short, well-validated measures for reading and mathematics. Future research could build on this study by relying on a more extensive assessment of school performance involving multiple measures. Second, we only collected reading fluency and mathematics data at the second time point but not at the first time point. An interesting question would be to investigate whether feedback learning and brain measures can predict reading and mathematics even better than tests for reading and mathematics themselves. On the other hand, an advantage of measuring feedback learning or other executive functioning tasks is that it captures abilities that are essential to both reading and mathematics. Third, IQ was assessed with only two subtests of the WISC/WAIS. A more comprehensive assessment of IQ might give a more definite answer to the question whether the relation between feedback learning and school performance is driven by underlying differences in general intelligence. Fourth, mathematics was assessed with the WISC for younger participants (10–15 years at T2) and with the WAIS for older participants (16–27 years at T2). When we performed the analyses with mathematics raw scores rather than norm scores (scores relative to same-aged peers), we needed to perform the analyses in separate age groups. These analyses showed that the prediction of mathematics scores from behavioral performance and neural activity for feedback learning was only present in the youngest age group (10–15 years). One tentative interpretation is that prediction is stronger in the younger age groups, when brain maturation is still undergoing major changes (Giedd & Rapoport, 2010). Alternatively, it is possible that the WISC scores are more sensitive for picking up change than the WAIS scores. Further studies should use a wider battery of tests to test these competing hypotheses in more detail.

It should also be noted that the working memory task we used in this study involved simple mental addition and subtraction. Therefore, this particular working memory task may be related more to mathematics than to reading performance. Also, younger children may be less proficient in these simple addition and subtraction skills, resulting in lower working memory outcomes. The working memory task contained both visuospatial elements as well as a verbal rehearsal component. In future studies, working memory could be controlled for with additional measures of working memory performance, to investigate the generalizability of these findings and the effects of visuospatial versus verbal working memory. In a similar way, the feedback learning task contains both visuospatial and verbal elements. Given that the correct stimulus needs to be sorted in one of three locations gives it a clear visuospatial aspect. However, many participants reported using a verbal rehearsal strategy during the task. Future studies should try to disentangle these components, given that they may contribute to reading and math in different ways.

Finally, in future research studying prediction of future behavioral outcomes from neuroimaging data, it is important to use statistical models that provide more definitive answers to the question whether prediction in an independent sample is possible. A recent review (Gabrieli, Ghosh, & Whitfield-Gabrieli, 2015) argues for the use of cross-validation analyses (such as the leave-

k-out method) to make claims about the prediction of behavior from neuromarkers stronger.

Conclusion

In conclusion, this study found contributions of feedback learning performance and neural activity in predicting school outcomes two years later. This provides evidence that studying learning processes through simplified laboratory tasks provides at least some relevance for real-world learning. In addition, we showed that neural measures explain unique variance in school outcomes 2 years later that is not captured by behavioral testing of feedback learning performance, working memory, or IQ alone. An important direction for future research is to unravel which predictors are involved in specific functions (such as reading or arithmetic) and which are involved in general school performance (e.g., predicting both reading and arithmetic). These results will be important to eventually tailor educational programs to the individual needs of children.

References

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology, 106,* 20–29. http://dx.doi.org/10.1016/j.jecp.2009.11.003

Arsalidou, M., & Taylor, M. J. (2011). Is 2 + 2 = 4? Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage, 54,* 2382–2393. http://dx.doi.org/10.1016/j.neuroimage.2010.10.009

Barceló, F., & Knight, R. T. (2002). Both random and perseverative errors underlie WCST deficits in prefrontal patients. *Neuropsychologia, 40,* 349–356. http://dx.doi.org/10.1016/S0028-3932(01)00110-5

Barnea-Goraly, N., Menon, V., Eckert, M., Tamm, L., Bammer, R., Karchemskiy, A., . . . Reiss, A. L. (2005). White matter development during childhood and adolescence: A cross-sectional diffusion tensor imaging study. *Cerebral Cortex, 15,* 1848–1854. http://dx.doi.org/10.1093/cercor/bhi062

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78,* 647–663. http://dx.doi.org/10.1111/j.1467-8624.2007.01019.x

Braams, B. R., van Duijvenvoorde, A. C. K., Peper, J. S., & Crone, E. A. (2015). Longitudinal changes in adolescent risk-taking: A comprehensive study of neural responses to rewards, pubertal development, and risk-taking behavior. *The Journal of Neuroscience, 35,* 7226–7238. http://dx.doi.org/10.1523/JNEUROSCI.4764-14.2015

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). Region of interest analysis using the MarsBar toolbox for SPM 99. *NeuroImage, 16,* S497.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19,* 273–293. http://dx.doi.org/10.1207/S15326942DN1903_3

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96,* 31–42. http://dx.doi.org/10.1037/0022-0663.96.1.31

Canty, A., & Ripley, B. (2016). *Boot: Bootstrap R (S-Plus) Functions*. R package version 1.3–18.

Cocosco, C. A., Kollokian, V., Kwan, R. K. S., & Evans, A. C. (1997). Online interface to a 3D MRI simulated brain database. *NeuroImage, 5,* S425.

Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency.

Trends in Neuroscience and Education, 3, 63–68. http://dx.doi.org/10.1016/j.tine.2013.12.001

Crone, E. A., Zanolie, K., Van Leijenhorst, L., Westenberg, P. M., & Rombouts, S. A. (2008). Neural mechanisms supporting flexible performance adjustment during development. *Cognitive, Affective & Behavioral Neuroscience, 8,* 165–177. http://dx.doi.org/10.3758/CABN.8.2.165

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping, 8,* 109–114. http://dx.doi.org/10.1002/(SICI)1097-0193(1999)8:2/3<109::AID-HBM7>3.0.CO;2-W

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64,* 135.

Dumontheil, I., & Klingberg, T. (2012). Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later. *Cerebral Cortex, 22,* 1078–1085. http://dx.doi.org/10.1093/cercor/bhr175

Eliassen, J. C., Lamy, M., Allendorfer, J. B., Boespflug, E., Bullard, D. P., Smith, M. S., . . . Strakowski, S. M. (2012). Selective role for striatal and prefrontal regions in processing first trial feedback during single-trial associative learning. *Brain Research, 1458,* 56–66. http://dx.doi.org/10.1016/j.brainres.2012.04.010

Ferstl, E. C., Neumann, J., Bogler, C., & von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping, 29,* 581–593. http://dx.doi.org/10.1002/hbm.20422

Fortier, M. S., Vallerand, R. J., & Guay, F. (1995). Academic motivation and school performance: Toward a structural model. *Contemporary Educational Psychology, 20,* 257–274. http://dx.doi.org/10.1006/ceps.1995.1017

Gabrieli, J. D. E., Ghosh, S. S., & Whitfield-Gabrieli, S. (2015). Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron, 85,* 11–26. http://dx.doi.org/10.1016/j.neuron.2014.10.047

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134,* 31–60. http://dx.doi.org/10.1037/0033-2909.134.1.31

Geary, D. C. (2000). From infancy to adulthood: The development of numerical abilities. *European Child & Adolescent Psychiatry, 9,* 11–16. http://dx.doi.org/10.1007/s007870070004

Giedd, J. N., & Rapoport, J. L. (2010). Structural MRI of pediatric brain development: What have we learned and where are we going? *Neuron, 67,* 728–734. http://dx.doi.org/10.1016/j.neuron.2010.08.040

Hoeft, F., Ueno, T., Reiss, A. L., Meyler, A., Whitfield-Gabrieli, S., Glover, G. H., . . . Gabrieli, J. D. (2007). Prediction of children's reading skills using behavioral, functional, and structural neuroimaging measures. *Behavioral Neuroscience, 121,* 602–613. http://dx.doi.org/10.1037/0735-7044.121.3.602

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109,* 679–709. http://dx.doi.org/10.1037/0033-295X.109.4.679

Houdé, O., Rossi, S., Lubin, A., & Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: An fMRI meta-analysis of 52 studies including 842 children. *Developmental Science, 13,* 876–885. http://dx.doi.org/10.1111/j.1467-7687.2009.00938.x

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44,* 2017–2036. http://dx.doi.org/10.1016/j.neuropsychologia.2006.01.010

Jenkins, J. R., Fuchs, L. S., Van Den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology, 95,* 719–729. http://dx.doi.org/10.1037/0022-0663.95.4.719

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80,* 437–447. http://dx.doi.org/10.1037/0022-0663.80.4.437

Krom, R., Jongen, I., Verhelst, N., Kamphuis, F., & Kleintjes, F. (2010). *DMT en AVI: Wetenschappelijke verantwoording* [DMT and AVI: Scientific basis]. CITO institute for Test Development.

Landi, N., Frost, S. J., Mencl, W. E., Sandak, R., & Pugh, K. R. (2013). Neurobiological bases of reading comprehension: Insights from neuroimaging studies of word level and text level processing in skilled and impaired readers. *Reading & Writing Quarterly, 29,* 145–167. http://dx .doi.org/10.1080/10573569.2013.758566

Li, Y., Hu, Y., Wang, Y., Weng, J., & Chen, F. (2013). Individual structural differences in left inferior parietal area are associated with school childrens' arithmetic scores. *Frontiers in Human Neuroscience, 7,* 844. http://dx.doi.org/10.3389/fnhum.2013.00844

Maurer, U., Bucher, K., Brem, S., Benz, R., Kranz, F., Schulz, E., . . . Brandeis, D. (2009). Neurophysiology in preschool improves behavioral prediction of reading ability throughout primary school. *Biological Psychiatry, 66,* 341–348. http://dx.doi.org/10.1016/j.biopsych.2009.02.031

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100. http://dx.doi.org/10.1006/ cogp.1999.0734

O'Shaughnessy, T. E., Lane, K. L., Gresham, F. M., & Beebe-Frankenberger, M. E. (2003). Children placed at risk for learning and behavioral difficulties implementing a school-wide system of early identification and intervention. *Remedial and Special Education, 24,* 27–35. http://dx.doi.org/10.1177/074193250302400103

Peters, S., Braams, B. R., Raijmakers, M. E. J., Koolschijn, P. C. M. P., & Crone, E. A. (2014). The neural coding of feedback learning across child and adolescent development. *Journal of Cognitive Neuroscience, 26,* 1705–1720. http://dx.doi.org/10.1162/jocn_a_00594

Peters, S., Koolschijn, P. C. M. P., Crone, E. A., Van Duijvenvoorde, A. C. K., & Raijmakers, M. E. J. (2014). Strategies influence neural activity for feedback learning across child and adolescent development. *Neuropsychologia, 62,* 365–374. http://dx.doi.org/10.1016/j.neuropsychologia.2014.07.006

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20,* 110–122. http://dx.doi.org/10.1016/j.lindif.2009.10.005

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *The Journal of Human Resources, 27,* 313–328. http://dx.doi.org/10.2307/145737

Somerville, L. H., Jones, R. M., Ruberry, E. J., Dyke, J. P., Glover, G., & Casey, B. J. (2013). The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychological Science, 24,* 1554–1562. http://dx.doi.org/10.1177/0956797613475633

St. Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 59,* 745–759. http://dx.doi.org/10.1080/ 17470210500162854

van den Bos, W., Güroğlu, B., van den Bulk, B. G., Rombouts, S. A., & Crone, E. A. (2009). Better than expected or as bad as you thought? The neurocognitive development of probabilistic feedback processing. *Frontiers in Human Neuroscience, 3,* 52. http://dx.doi.org/10.3389/neuro.09 .052.2009

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of Experimental Child Psychology, 87,* 239–266. http://dx.doi .org/10.1016/j.jecp.2003.12.002

van Duijvenvoorde, A. C. K., Zanolie, K., Rombouts, S. A., Raijmakers, M. E., & Crone, E. A. (2008). Evaluating the negative or valuing the positive? Neural mechanisms supporting feedback-based learning across development. *The Journal of Neuroscience, 28,* 9495–9503. http://dx.doi .org/10.1523/JNEUROSCI.1485-08.2008

Zanolie, K., Van Leijenhorst, L., Rombouts, S. A., & Crone, E. A. (2008). Separable neural mechanisms contribute to feedback processing in a rule-learning task. *Neuropsychologia, 46,* 117–126. http://dx.doi.org/10 .1016/j.neuropsychologia.2007.08.009