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

Peters, S.

Citation

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

Author: Peters, Sabine
Title: The adolescent brain : unraveling the neural mechanisms of cognitive and affective development
Issue Date: 2016-01-27

Chapter 3

Neural activity for feedback learning in adults

This chapter is based on:

Peters, S., Koolschijn, P.M.C.P., Raijmakers, M.E.J., Schrijver, M.S., Overgaauw, S. & Crone, E.A. Feedback learning: Neural reactions to valence and informative value depend on individual differences in strategy use (submitted for publication, 2014).

Abstract

Learning from feedback is an important aspect of flexible behavior, but how neural regions contribute to complex feedback learning is not yet understood. Previous studies found that a widespread brain network is activated for feedback learning, including the dorsolateral prefrontal cortex (DLPFC), pre-supplementary motor area/anterior cingulate cortex (pre-SMA/ACC) and superior parietal cortex (SPC). In this fMRI study, participants performed a multiple-trial learning task where a distinction was made between a rule learning phase and a rule application phase, and between positive and negative learning, so that we could investigate which areas are sensitive to learning value and valence of feedback. We studied feedback processing in a more complex learning environment, where both positive and negative feedback were similarly informative for learning. This way, we addressed the issue that in previous studies, positive and negative feedback often differed in informative value. Additionally, finite mixture modeling on task performance allowed us to distinguish two different latent learning strategies, which allowed us to further pinpoint individual differences in neural activation. The results indicated that the DLPFC, pre-SMA and SPC were all sensitive to learning value of feedback (feedback during the learning phase compared to feedback during the application phase), with all regions, except for left SPC, showing more activation after negative learning compared to positive learning. Additionally, individual differences in strategy use were related to the degree to which areas in the feedback learning network were recruited, such that high performers showed faster learning and increased neural activation to learning signals in general. Together, these findings highlight the need for a more mechanistic understanding of complex learning, for which a modeling approach of strategy types proved to be specifically valuable.

Introduction

One of the key elements of successful learning is the ability to use performance feedback in order to adjust future behavior. Feedback can be either positive, signaling a continuation of current behavior, or negative, indicating the need for a behavioral adjustment (Holroyd & Coles, 2002). Learning from both positive and negative feedback signals is essential for successful adaptation to a changing environment. Even though much progress is made in understanding neural responses to single-trial positive and negative feedback, much less is known about how these neural regions respond to learning in a complex environment across multiple trials. In this study, we investigated the neural basis of positive and negative feedback learning in a complex repeated learning task. Second, we employed hidden Markov models to categorize learners in strategy types, in order to relate neural activation after feedback to latent strategy differences.

Previous research has indicated that a widespread network of brain regions is involved in feedback processing, including the dorsal anterior cingulate cortex (dACC)/pre-supplementary motor area (pre-SMA) (Holroyd et al., 2004; Mars et al., 2005; Monchi, Petrides, Petre, Worsley, & Dagher, 2001; Ullsperger & von Cramon, 2003), (dorso)lateral prefrontal cortex (DLPFC) (Dove, Pollmann, Schubert, Wiggins, & von Cramon, 2000; Konishi et al., 2002; Lie, Specht, Marshall, & Fink, 2006; van den Bos et al., 2009; van Duijvenvoorde et al., 2008; van Veen et al., 2004; Zanolie et al., 2008), caudate nucleus (Monchi et al., 2001; Tricomi, Delgado, McCandliss, McClelland, & Fiez, 2006) and superior parietal cortex (SPC) (van Duijvenvoorde et al., 2008; Zanolie et al., 2008). Many of these studies reported more activation after negative feedback compared to positive feedback. For example, a paradigm that has frequently been used to investigate performance monitoring and feedback learning across multiple trials is the Wisconsin Card Sorting Task (WCST; Milner, 1963). In the WCST, participants are instructed to sort cards according to different sorting rules (based on number, color or shape), followed by positive or negative performance feedback. In addition, the task encompasses occasional unexpected rule switches, signaled by negative feedback. These studies, and related switching studies, reported specifically more DLPFC and pre-SMA/ACC activity after negative than positive feedback (Dove et al., 2000; Konishi et al., 2002; Lie et al., 2006; Zanolie et al., 2008). Together, these studies led to the conclusion that this network is important for expectation violation and performance adjustment.

Despite these consistent findings, a second set of studies has reported more DLPFC activity after positive feedback (van den Bos et al., 2009; van Veen et al., 2004), which does not fit well with the hypothesis of either expectation violation or performance adjustment. For example, in a probabilistic learning task, the DLPFC and SPC were more active after positive feedback when participants were applying the correct rule, compared to negative feedback (van den Bos et al., 2009). One proposed explanation for these conflicting findings is that the DLPFC and SPC respond to informative value of feedback rather than valence. In probabilistic designs, negative feedback after applying the correct rule does not indicate that a switch in behavior is necessary. The previous experimental paradigms thus differed in the informative value that was provided by positive and negative feedback; in the latter probabilistic paradigm (van den Bos et al., 2009) positive feedback could be more informative than negative feedback. One way to unravel these inconsistent findings is by studying feedback processing in a more complex but deterministic learning environment, in which both positive and negative feedback are important for learning but might differ in informative value between trials.

Prior studies have shown that collapsing performance and using simple performance indices as correlates for neural activation fails to take into account much of the variance in behavior. In previous behavioral studies it was found that individual differences in learning performance are partially due to categorically different strategies that participants apply (Johansen & Palmeri, 2002; Schmittmann, Visser, & Raijmakers, 2006; Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003). These strategies can be characterized based on the degree to which participants take into account the informative value of feedback. These prior studies hypothesized that neural engagement differs depending on the efficiency of strategy use (Schmittmann, van der Maas, & Raijmakers, 2012), and having a better index of these strategy groups may allow us to further understand how neural activation in DLPFC, pre-SMA and SPC is important for complex learning. Thus, there is a high need for a more mechanistic understanding of how individuals learn in complex environments; therefore, we made use of a modeling approach to dissociate low-strategy and high-strategy learners.

In the current study, we used a higher-order rule-learning paradigm, which distinguished different feedback types based on early and late learning. This idea relates to prior research in which early and late learning were compared (Brovelli, Laksiri, Nazarian, Meunier, & Boussaoud, 2008; Eliassen et al., 2012). For example, Eliassen et al. (2012) investigated which neural areas are sensitive to the learning value of feedback by comparing activity during feedback on first trials of a learning sequence with feedback during later trials. Results showed that several brain regions were uniquely activated by first trial feedback, including the striatum, DLPFC and parietal and temporal areas. Here, we developed a learning task in which participants had to categorize stimuli in one of three locations. This paradigm allowed us to study neural responses to the informative value of positive and negative feedback in a learning phase of the task. In addition, it allowed us to model behavior according to different strategies.

First, we predicted that neural responses in DLPFC, SPC, and pre-SMA/ACC to both positive and negative feedback would be larger when learning compared to when applying correct sorting rules. This hypothesis was based on a prior study that found that for novel higher-order rule-learning, the DLPFC and pre-SMA were more active for early than late learning, consistent with the idea that these areas are sensitive to informative value for learning (Boettiger & D'Esposito, 2005). Second, we addressed the question whether negative feedback would result in additional activation in DLPFC and pre-SMA/ACC during learning compared to positive feedback, because of their presumed role in expectation violation and performance adjustment (Holroyd et al., 2004). Third, we hypothesized that low-strategy learners would be less sensitive to the informative value of feedback, leading to less activation in the feedback learning network during learning, when compared to high-strategy learners.

Methods

Participants

Thirty-two healthy young adults were included in the study (16 women, 18-25 years old, *M* = 21.39, *SD* = 1.88). Participants were paid volunteers, recruited through local advertisements or through a university course credits system. All participants were right-handed and reported no history of neurological or psychiatric disorders or (past or current) use of psychotropic medication. Informed consent was obtained and the study was approved by the Internal Review Board at the Leiden University Medical Center. All anatomical scans were reviewed and cleared by a radiologist. One female participant (not included in the demographics above) was excluded from the analysis due to excessive movement in the MRI scanner (> 5mm). IQ was estimated with two subtests of the WAIS-IIIR (Similarities and Block Design). All estimated IQ-scores were within the normal range (*M* = 113.05, *SD* = 9.13, range = 100–130).

Feedback Learning Task

Participants performed a feedback learning task in the MRI scanner. They were presented with three empty boxes under which three possible stimuli were presented one by one, for multiple times in a random order (Figure 1).

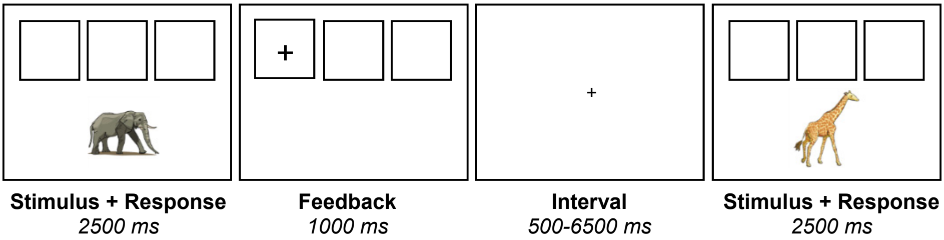

Figure 1: Sorting task, display of task sequence. Participants sorted stimuli on one of three locations and used performance feedback to find the correct location for three stimuli that all correspond to one of the boxes.

The participants were informed that each of the three stimuli belonged in one of the three boxes. By trial-and-error and using performance feedback, the correct location for all three stimuli could be found. Performance feedback was presented in the form of a plus-sign for positive feedback and a minus-sign for negative feedback. After applying the correct location twice for each stimulus, or after twelve trials, a new sequence with new stimuli was presented. New stimuli were chosen to prevent participants from trying to find a relation between sequences. In total, fifteen sequences were presented, which resulted in a maximum of 180 trials. Prior to scanning, participants practiced three sequences similar to the experimental task to ensure they understood the task. The task was divided in two blocks of eight and seven sequences each. The stimuli were presented in a pseudorandom order, with a maximum of two identical stimuli in a row. Stimuli were 250x250 pixels clip-art images found through Google Image search. Interstimulus intervals were jittered using OptSeq (Dale, 1999; see also http://surfer.nmr.mgh.harvard.edu/optseq), with intervals varying between 0-6 seconds.

Feedback types

For each of the three stimuli, a distinction was made between trials in the learning phase and the application phase. The learning phase was defined as those trials when participants had not yet found the correct location of the stimulus, and were guessing or reasoning to find the correct location. The application phase was when each stimulus was already sorted correctly at least once. Feedback in the learning phase was only scored as 'learning' based on subsequent responses for the same stimulus, i.e. choosing the same response (after positive feedback) or choosing a different response (after negative feedback). We were only interested in trials that actually resulted in learning; therefore trials in the learning phase that did not result in learning were excluded from further analysis (*M* = 4.2%, *SD* = 3.7% of the trials). The four feedback types were defined as follows:

Learning phase. (a) Positive learning: The first correct response for a stimulus, if followed by a correct response the next time that stimulus appears. (b) Negative learning: An incorrect response for a stimulus that had not yet been sorted correctly. The next time that stimulus appeared, the participant did not perseverate but chose another option.

Application phase. Choosing the correct location for a stimulus that was sorted correctly before.

Learning performance

Learning rate was calculated for each participant as a measure of learning performance. This was defined as the percentage of trials in the learning phase, where feedback was successfully used on the next trial.

Strategy performance

A second way in which performance relations were studied was by looking at the efficiency of strategy use. In the experimental task, several strategies differing in efficiency could be employed to learn the correct stimulus-location relations. The most optimal strategy accounts not only for feedback on the former responses of the same stimulus but also for other stimuli. This is based on two insights into the learning task: first, a location can only be associated with one stimulus (hence, choose a location that is not already correctly associated with another stimulus), and second, the probability of a correct response depends on the number of stimulus candidates for the chosen location (hence, choose a location for which negative feedback was given with another stimulus).

For each trial the efficiency of a response was defined as follows, in ascending order of the complexity of reasoning that is violated (note that for strategy analysis the non-learning trials were included): (a) Mistake: repeating a previously made error (no reasoning required to know this is an incorrect response) or making an error after an earlier correct response for the same stimulus. (b) Inefficient: when the location of a picture is known, the participant fails to deduce this location is not the correct location of another picture. (c) Suboptimal: when participants learned that a picture does not belong in a certain location (and if that is the only knowledge), the optimal decision is to place another stimulus in that location, to ensure a 50 percent chance of being correct. Failing to use this strategy is scored as a suboptimal decision. (d) Optimal: all other cases.

Strategy groups

We expected latent groups of participants that employed different learning strategies differing in the complexity of reasoning. We aimed to test: (a) how many strategies we could distinguish, and (b) which strategy each participant employed. Based on time series of trials that were coded in terms of the efficiency of responses (4-valued multinomial data), latent strategy performance groups were determined using finite mixtures of log-linear regression models (Leisch, 2004; McLachlan & Peel, 2000). More advanced hidden Markov models were applied as well (Visser, 2011), but these analyses resulted in less optimal models. Finite-mixture models assume that there are discrete components each of which produces series of responses according to a dedicated probability distribution, which characterizes a learning strategy. In this case, each component is a log-linear regression model, with trial number of the sub series as a possible predictor variable. The responses of each participant are assumed to be generated according to one component, but component membership is a latent variable, i.e., it cannot be observed directly from the data. To fit a mixture model to data, the number of components must be specified. The problem of selecting this parameter is known as 'model selection'. It can be resolved by using the minimum of the Bayesian information criterion (BIC), which implements a trade-off between log-likelihood of the model and the number of free parameters. The model with the lowest BIC is the optimal model.

We fitted mixture models with one, two or three components to our data; components were defined with and without trial number as a predictor variable. After establishing the optimal model, we assigned each individual to the component for which his/her data was most likely, according to the posterior probabilities of the data given the model (Visser, 2011). We used the depmixS4 software (Visser & Speekenbrink, 2010) for the subsequent analyses.

FMRI Data Acquisition

Scans were acquired with a standard whole-head coil on a Philips 3.0 Tesla MRI scanner at the Leiden University Medical Center. The functional scans were acquired using a T2*-weighted echo-planar imaging (EPI) during two functional runs with a variable number of volumes per subject, because the length of the task varied per subject. The first two volumes were discarded to allow for equilibration of T1 saturation effects (TR = 2.2 sec, TE = 30 ms, sequential acquisition, 38 slices of 2.75 mm, field of view 220 mm, 80x80 matrix, in-plane resolution 2.75 mm). A high-resolution 3D T1-FFE scan for anatomical reference was obtained (TR = 9.760 ms; TE = 4.59 ms, flip angle = 8 degrees, 140 slices, $0.875 \times 0.875 \times 1.2$ mm^3 voxels, FOV = $224 \times 168 \times 177$ mm^3). Head motion was restricted using a pillow and foam inserts that surrounded the head. The experimental task was projected on a screen that was viewed through a mirror.

FMRI Data Analysis

All data was analyzed with SPM8 (Wellcome Department of Cognitive Neurology, London). Images were corrected for differences in timing of slice acquisition and rigid body motion. Structural and functional volumes were spatially normalized to T1 templates. Translational movement parameters never exceeded 1 voxel (< 3 mm) in any direction for any participant or scan. The normalization algorithm used a 12-parameter affine transformation together with a nonlinear transformation involving cosine basis functions and resampled the volumes to 3 mm cubic voxels. Templates were based on the MNI305 stereotaxic space (Cocosco, Kollokian, Kwan, & Evans, 1997), an approximation of Talairach space (Talairach & Tourneaux, 1988). Functional volumes were spatially smoothed with an 8mm FWHM isotropic Gaussian kernel. Statistical analyses were performed on individual participant's data using the general linear model in SPM8. The fMRI time series data were modeled by a series of events convolved with a canonical hemodynamic response function. The feedback of each trial was modeled as an event of interest. The trial functions were used as covariates in a general linear model; along with a basic set of cosine functions that high-pass filtered the data, and a covariate for session effects. The least-squares parameter estimates of height of the best-fitting canonical HRF for each condition were used in pair-wise contrasts. The resulting contrast images, computed on a subject-by-subject basis, were submitted to group analyses. Task related responses were considered significant if they consisted of at least 10 contiguous voxels that exceeded a stringent threshold $p < .05$ (FWE-corrected).

Region-of-interest Analysis

Region-of-interest (ROI) analyses were performed with the Marsbar toolbox in SPM8 (Brett, Anton, Valabregue, & Poline, 2002). ROIs that spanned several functional brain regions (in our case, all ROIs) were subdivided by sequentially masking the functional ROI with each of several anatomical Marsbar ROIs. The contrast used to generate functional ROIs was based on the general contrast activation > fixation, (FDR-corrected, $p < .05$, 10 contiguous voxels) across all participants. For all ROI analyses, effects were considered significant at $\alpha = 0.008$ based on Bonferroni correction for six ROIs (unless reported otherwise).

Results

Behavioral results

Performance on the task was sufficient for all participants. On average, participants needed 135 (range 125-147, *SD* = 5.94) trials to complete the task (out of a maximum number of 180 trials). In general, participants had a high learning rate (*M* = 95.82%, *SD* = 3.71%).

Similar to learning rate, strategy performance was high (*M* = 94.84%, *SD* = 4.64% optimal trials). Strategy performance correlated strongly with learning rate ($r = 0.86$, $p < .001$). Suboptimal decisions (*M* = 1.57%, *SD* = 2.17%), inefficient decisions (*M* = 2.14%, *SD* = 1.81%) and mistakes (*M* = 1.45%, *SD* = 1.53%) were relatively rare. No sex differences were found for learning rate or strategy performance.

Strategy performance group

The optimal mixture model, i.e. the model with the lowest BIC, was one with two performance groups, referred to as the 'low-strategy learners' and the 'high-strategy learners' (see Table 1 for the fit statistics of the fitted models). The high-strategy learners have stable response probabilities over sub trials with very high probability of optimal responses. The low-strategy-learners have response probabilities that depend on sub trial number, such that suboptimal and inefficient responses mainly occur during the first few trials. Figure 2 shows the observed response probabilities for each performance group. The two groups did not differ in age, sex, or estimated IQ. As expected, low-strategy and high-strategy performers differed in learning rate ($t = 4.60$, $p < .001$), such that high-strategy learners had a higher learning rate than low-strategy learners, for both positive learning ($t = 2.54$, $p = .017$) and negative learning ($t = 4.68$, $p = .001$). To investigate possible differences in the relative contributions of positive and negative learning to the overall learning rate, the ratio of positive learning compared to negative learning rate was also calculated. This analysis showed that low-strategy performers learned relatively more from positive feedback than high-strategy performers ($t = 2.99$, $p = .014$) and high-strategy performers learned relatively more from negative feedback than low-strategy performers ($t = 3.59$, $p = .001$).

Table 1: Fit statistics of latent logistic regression models. The number of observations is 4297. Loglike is the Log-likelihood of the model; df is the number of freely estimated parameters in the model; BIC is the Bayesian Information Criterion. Models consist of 1, 2, or 3 components. Each component is a logistic regression function with an intercept (1) or an intercept + trial number as a predictor variable (1b). The asterisk indicates the parsimonious, best fitting model (with the lowest BIC).

Model	loglike	df	BIC
1	-1134.5	3	2294.1
1b	-1073.7	6	2197.6
1, 1	-1094.7	7	2248.0
1, 1b*	-1041.6	10	2166.8
1b, 1b	-1032.9	13	2174.5
1, 1, 1	-1088.7	11	2269.4
1, 1, 1b	-1041.4	14	2200.0

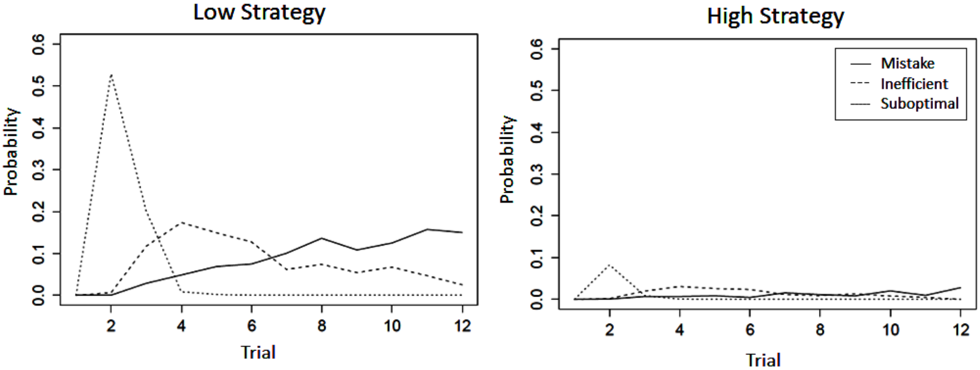

Figure 2: These figures display the probability of making a response at each trial for Low-strategy participants (left figure) and High-strategy participants (right figure). The different lines indicate different types of trials: mistakes, inefficient trials, sub-optimal trials. The remaining type of trials (optimal trials), are not shown.

Whole brain analyses

Learning versus application

To test which areas were responsive to learning value, we focused on the contrast learning (both positive and negative) > application. The regions that were more active for the learning > application contrast are presented in Figure 3, and included bilateral lateral PFC, pre-SMA and bilateral parietal cortex.

Positive learning vs. application

To test the hypothesis that the feedback learning network can also be activated by positive feedback, provided it is high in informative value for learning, the contrast positive learning > application was calculated. In both conditions positive feedback was provided, which ensures the only difference between conditions is informative value for learning. This contrast also resulted in robust activation in bilateral lateral PFC, pre-SMA and bilateral parietal cortex. As can be seen in Figure 3, there was high overlap in brain regions that were recruited for positive learning vs. application, and negative learning vs. application (see overlap in Figure 3). In the next section, positive and negative learning are tested against each other.

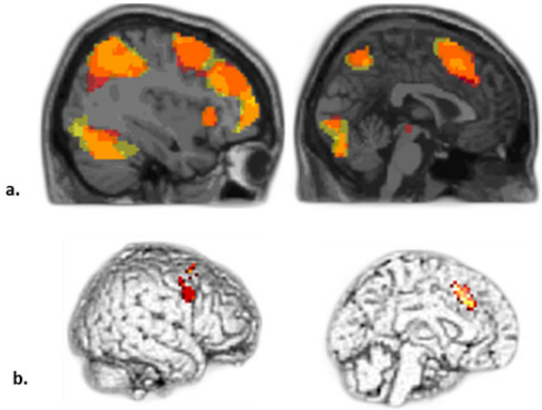

Figure 3: (a): Negative learning compared to application (red), positive learning compared to application (yellow), and overlap (orange). Results are FWE-corrected, p < .05, 10 contiguous voxels. The left image shows a lateral slice (x: 37 y: 8 z: 32) and the right image a medial slice (x: 1 y: 8 z: 12). (b): Negative learning > positive learning, FWE-corrected, p < .05, 10 contiguous voxels.

Positive learning vs. negative learning

To test which areas were sensitive to feedback valence, we focused on the contrast negative learning > positive learning and vice versa. We chose to compare positive and negative learning (instead of all negative > all positive feedback) to ensure that the only difference between the conditions was valence, and value for learning was comparable. The results demonstrated more activity for negative learning compared to positive learning in the right DLPFC, pre-SMA, anterior and middle cingulum. The reverse contrast, positive compared to negative learning resulted in activation in the bilateral supramarginal gyrus and bilateral superior temporal cortex (see Table 2).

Table 2: Areas that are activated for the contrast negative feedback > positive feedback during learning, FWE-corrected, 10 contiguous voxels.

Area of activation	**x**	**y**	**z**
R superior medial frontal gyrus	9	21	42
L superior medial frontal gyrus	-6	30	33
R superior frontal gyrus	21	3	54
R opercular inferior frontal gyrus	42	9	36
R precentral gyrus	36	0	48

Learning vs. application in low-strategy and high-strategy learners

Neural differences between high-strategy and low-strategy learners were investigated with a whole brain contrast comparing learning > application in the two groups. Uncorrected results are reported because the relatively small group of low-strategy learners did not survive FWE or FDR correction. A two-sample *t*-test (uncorrected, $p < .001$, 10 contiguous voxels) revealed that high-strategy learners showed more activation in right DLPFC and pre-SMA compared to low-strategy learners for the contrast learning > application (see Figure 4, and Table 3).

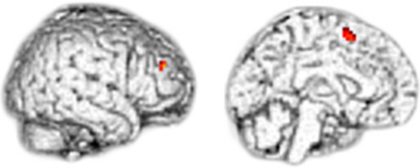

Figure 4: Areas that were more active for the contrast learning > application in high-strategy learners compared to low-strategy learners, $p < .001$, uncorrected, 10 contiguous voxels.

Table 3: Areas that are more activated for the contrast learning > application in high-strategy learners compared to low-strategy learners, p < .001, uncorrected, 10 contiguous voxels.

Area of activation	**x**	**y**	**z**
L opercular inferior frontal gyrus	-45	9	12
L opercular inferior frontal gyrus	-57	6	6
L inferior temporal gyrus	-39	-21	-21
L opercular inferior frontal gyrus	-33	6	21
L supplementary motor area	-3	12	57
R middle frontal gyrus	33	42	27

Region of Interest (ROI) analysis

The analyses above illustrate that a network of areas is involved in feedback learning. To further characterize sensitivity to valence and to investigate relations with performance, we performed a ROI analysis on five a priori defined regions based on the contrast all conditions > fixation. Two separate regions were found in the right DLPFC, resulting in a total of six ROIs (left DLPFC (center-of-mass: x: -38, y: 14, z: 48), right DLPFC (x: 44, y: 34, z: 32), right superior DLPFC (x: 35, y: 10, z: 55), pre-SMA (x: -8, y: 11, z: 54), left SPC (x: -30, y: -64, z: 57) and right SPC (x: 32, y: -62, z: 57). All significant results survived Bonferroni correction unless otherwise specified. Figure 5 shows a representative subset of the results.

Areas sensitive to learning value

All areas that were included in the ROI analysis were sensitive to learning value. This was defined as relatively more activity for feedback during learning than during application. That is to say, there were main effects of feedback type for all six ROIs (all *p*s < .001). Follow up paired-samples t-test comparisons indicated that all ROIs were more active during positive learning relative to application (all *p*s < .001) and negative learning relative to application (all *p*s < .001).

Areas sensitive to feedback valence

Next, we compared negative learning and positive learning to investigate which areas were sensitive to feedback valence in the learning phase.

- *DLPFC.* Right DLPFC and right superior DLPFC were both more active after negative learning than positive learning (right DLPFC: $t(31) = 4.57$, $p < .001$, right superior DLPFC ($t(31) = 4.51$, $p < .001$). The left DLPFC revealed a similar pattern ($t(31) = 2.73$ $p = .010$[1]). A direct comparison of valence (positive and negative learning) for the left and right DLPFC revealed that the difference between negative learning and positive learning was larger in right than in left DLPFC (Interaction Region x Condition; $F(1,31) = 10.41$, $p = .003$).

---

[1] This effect does not survive Bonferroni correction

- *Pre-SMA.* The pre-SMA was more active after negative learning than positive learning ($t(31) = 4.26$, $p < .001$). A direct comparison of region and valence showed that the right DLPFC and pre-SMA has a similar pattern for positive and negative learning (Interaction Region x Condition; $F(1,31) = .32$, $p = .58$). The left DLPFC and pre-SMA, however, differed in neural activity pattern ($F(1.31) = 16.59$, $p < .001$), such that the difference between negative learning and positive learning was larger in pre-SMA than in left DLPFC.

- *SPC.* The right SPC was more active after negative than after positive learning ($t(31) = 3.29$, $p = .003$). Left SPC did not differentiate between positive and negative learning ($t(31) = 1.44$, $p = .160$) A comparison of region and valence confirmed that left and right SPC demonstrate a different pattern (Interaction Region x Condition; $F(1,31) = 8.65$, $p = .006$).

Taken together, all a priori selected ROIs were more active for positive and negative learning compared to application, and all ROIS, except for left SPC (and left DLPFC without multiple comparisons corrections) were additionally more sensitive to negative learning than to positive learning.

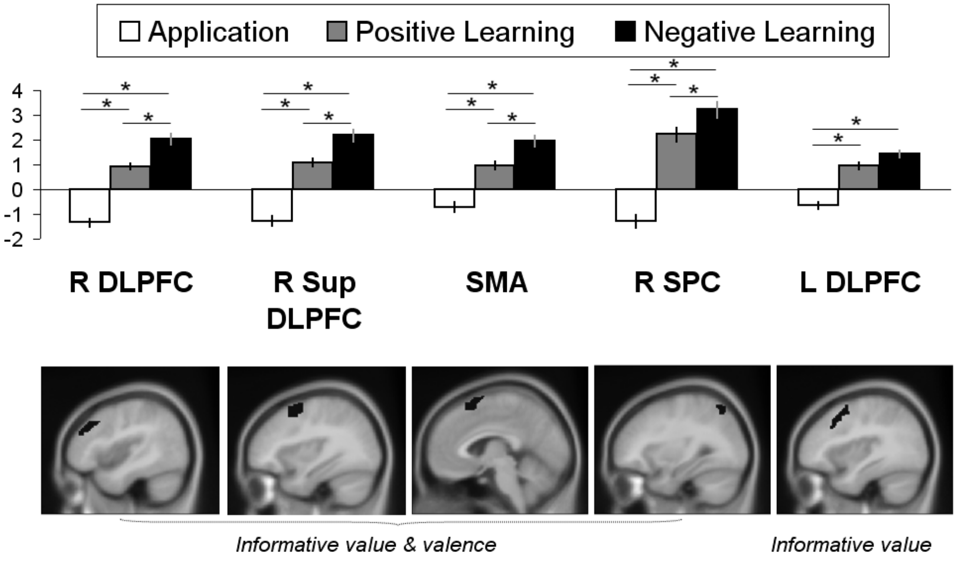

Figure 5. Region of Interest analysis Activation for application, positive learning, and negative learning. Abbreviations: L = left, R = right, DLPFC = dorsolateral prefrontal cortex, SMA = supplementary motor area, SPC = superior parietal cortex.

Correlations between neural activity and performance

Neural activity in the right superior DLPFC for learning > application was positively associated with learning rate in the task (r =.49, p = .005). A similar correlation was found for right SPC (r = .50, p = .004). Other regions did not show correlations with learning rate. For the strategy performance measures, we found that neural activation for learning > application correlated negatively with the percentage of mistakes in right SPC (r = -.36, p = .040[1]); and with the percentage of inefficient decisions in right superior DLPFC (r = -.46, p = 0.008), left DLPFC (r = -.37, p = .037[1]) and right SPC (r = -.49, p = .004). Taken together, right DLPFC and right SPC showed the most consistent relations to performance, but detailed strategy analyses confirmed that left DLPFC also contributed to efficient decisions in the task.

Discussion

Three main findings emerged from this study: (1) core areas in the feedback learning network (bilateral DLPFC, bilateral SPC and pre-SMA) were all sensitive to the learning value of feedback. Importantly, this effect was also found for positive feedback, but only when the positive feedback was informative for learning, (2) bilateral DLPFC, right SPC and pre-SMA were additionally more sensitive to negative feedback than positive feedback during learning, and (3) the degree to which areas in the feedback learning network were activated was dependent on learning performance and strategy use.

By comparing feedback during learning with feedback during application, it was found that a widespread network including DLPFC, pre-SMA and SPC was more active during learning. Our results are consistent with a prior single-trial learning study demonstrating a widespread network including the DLPFC and parietal areas which was sensitive to learning value (Eliassen et al., 2012). The current results also fit well with a prior study by Boettiger and D’Esposito (2005), which showed that complex learning of abstract stimuli that cannot easily be verbalized is associated with activation in the pre-SMA and DLPFC.

In prior studies, experimental designs were often developed such that positive feedback also revealed the rules for other stimuli, and with working memory demands that were often low (Dove et al., 2000; Lie et al., 2006; Zanolie et al., 2008). This study shows that positive feedback can result in similar neural activation as negative feedback when it is presented in a more complex learning setting that requires multiple trial learning. Prior studies have remained inconclusive about whether areas in the feedback learning network respond more to negative than to positive feedback. With stringent Bonferroni correction for multiple comparisons, three areas were found to be more active after negative feedback than positive feedback during learning (right DLPFC, pre-SMA and right SPC). This effect was also found in the left DLPFC, but only without multiple comparisons corrections.

The finding that the pre-SMA is more active after negative feedback is consistent with a number of other studies that have reported increased pre-SMA activation after negative feedback (Ozyurt, Rietze, & Thiel, 2012; Volz, Schubotz, & von Cramon, 2005). Second, the increased activation for negative versus positive feedback was larger for right DLPFC than for left DLPFC. These findings fit well with prior studies, in which it was suggested that the right DLPFC (together with the pre-SMA/ACC) is more important for performance adjustments than left DLPFC (Kerns et al., 2004), whereas the left DLPFC may be relatively more important for verbal working memory (Smith & Jonides, 1999; Wager & Smith, 2003). Notably, only right but not left SPFC showed sensitivity to valence. These findings suggest that the right hemisphere network in general is more important for performance adjustment.

The third question that was addressed in this study concerned individual differences in neural activity and learning performance. In the first set of general analyses, we found that learning rate was positively correlated with differential activation for learning compared to application in right DLPFC and, less strongly, in right SPC. Similar brain-behavior correlations in right DLPFC and right SPC were reported previously (Boettiger & D'Esposito, 2005; Koolschijn et al., 2011), and have been associated with a need for performance adjustment.

More importantly, we investigated whether differences in neural activity could be related to a categorization of strategy use. This method is valuable because it may detect latent differences in problem solving styles that cannot always be identified based on straightforward indices of behavior (Schmittmann et al., 2012). Indeed, participants differed in the strategies they employed during learning, and two different groups could be distinguished on the basis of strategy use. It was found that high-strategy learners showed more activity on a whole brain comparison of learning versus application in bilateral DLPFC and pre-SMA. Similarly, brain-behavior correlations were found for strategy performance in several regions within left and right DLPFC, demonstrating that especially these areas are related to individual differences in learning strategy (Andersen et al., 2014). Future studies should focus on deviant learning groups, and investigate whether the combined neuroimaging-behavioral modeling approach allows for the understanding of more variance in strategies.

In conclusion, with this study it was demonstrated that for complex feedback learning by reasoning, a network including DLPFC, pre-SMA and SPC is recruited that is sensitive to the learning value of both positive and negative feedback. Furthermore, all these regions, except for left SPFC, were additionally sensitive to valence. Detailed strategy categorizations confirmed that especially left and right DLPFC are associated with strategy differences in learning. These findings highlight the importance of distinguishing different subtypes of positive and negative feedback based on value for learning. In future research, the transition towards studying individual differences in strategy use and studying types of learning that more closely resemble real-life learning provides a promising way towards unraveling neural correlates for latent constructs of learning.