Individual differences in behavioral flexibility in a probabilistic reversal learning task : an fMRI study

Reference:

Waegeman Anja, Declerck Carolyn, Boone Christophe, Seurinck Ruth, Parizel Paul M..- *Individual differences in behavioral flexibility in a probabilistic reversal learning task : an fMRI study*

Journal of neuroscience, psychology, and economics - ISSN 2151-318X - 7:4(2014), p. 203-218

DOI: http://dx.doi.org/doi:10.1037/npe0000026

Title page with all author information

Individual differences in behavioural flexibility in a probabilistic reversal learning task: an fMRI study

Anja Waegeman[1,2], Carolyn H. Declerck[2], Christophe Boone[2], Ruth Seurinck[3] and Paul M. Parizel[4]

[2]University of Antwerp, Faculty of Applied Economics, Department of Management
B-2000 Antwerp, Belgium

[3] University of Ghent, Faculty of Psychology and Educational Sciences, Department of Data Analysis
B-9000 Ghent, Belgium

[4] University of Antwerp, Faculty of Medicine, Department of Radiology
B-2650 Edegem, Belgium

[1] Corresponding author at the Faculty of Applied Economics, Department of Management,
University of Antwerp, Prinsstraat 13, B -2000 Antwerp, Belgium
e-mail: anja.waegeman@uantwerp.be
phone : 0032 3 265 51 37
fax : 0032 3 275 50 79

Individual differences in behavioural flexibility in a probabilistic reversal learning task: an fMRI study

Submitted 27/01/2014

Revision 22/05/2014

2nd revision 25/09/2014

Behavioural flexibility is an important aspect of self-regulation and involves effectively learning, unlearning, and relearning associations between actions and outcome. Using a probabilistic reversal learning paradigm (PRL), the neural correlates of flexibility have previously been associated with brain regions implicated in cognitive control, including the anterior cingulate cortex (ACC) and lateral prefrontal cortex, and with the nucleus accumbens (Nacc) implicated in reward. The current study on healthy young males (N = 40) extends this previously published work in three ways. First we corroborate the involvement of ACC, VLPFC and DLPFC at the exact moment of behavioural switches. Second, we report increased activation of the dACC and caudate head with increasing number of perseverating errors preceding a behavioural switch. Third, better performance on the task is associated with increased activation of rACC and VLPFC during switching, suggesting that these regions contribute to individual differences in behavioural flexibility. These findings cannot be extended to individual differences in a self-reported measure of self-regulation.

Keywords: self-regulation, behavioural flexibility, probabilistic reversal learning, caudate nucleus, fMRI

1. Introduction

When a goal is set, impulse control and persistence to overcome the obstacles that hamper achievement of the goal is only part of the success. In a fast changing environment, we must be sure that the initial step remains the best cause of action to reach the goal. When the context in which our actions take place changes we should notice the change in action-outcome contingencies and adapt our behaviour based on the new knowledge. This ability is known as behavioural flexibility and, together with other executive functions such as impulse control and persistence, they form the basic components of human self-regulation. Behavioural flexibility allows people to maintain a balance between exploration and exploitation, holding on to a certain behaviour as long as it is adaptive, and adjust it in favour of a new strategy when it loses saliency (Cohen, McClure, & Yu, 2007). Because behavioural flexibility is a two-step process – giving up old habits and relearning new ones - it requires more cognitive capacity than mere learning through feedback.

Not all individuals show sufficient cognitive capacity to react in an equally flexible manner to environmental changes (Smillie, Cooper, Tharp, & Pelling, 2009). Impaired behavioural flexibility is often associated with clinically diagnosed pathological symptoms such as deficient executive functions, undirected repetitive thoughts, and rigid behaviours as observed in obsessive compulsive disorder (OCD). The behavioural rigidity of this disorder has been associated with altered functioning of several brain regions in the frontostriatal circuit (Aouizerate et al., 2004) such as reduced activation of the orbitofrontal cortex (OFC) (Chamberlain et al., 2008) and hyperactivity of the head of the caudate nucleus (Guehl et al., 2008; Whiteside, Port, & Abramowitz, 2004). Interestingly, despite an overactive caudate nucleus, the functional connectivity between the caudate head and the dorsolateral prefrontal cortex (DLPFC) is reduced in obsessive compulsive disorder (Harrison et al., 2009). Via the frontostriatal pathway connecting the DLPFC with the caudate nucleus (Tekin & Cummings, 2002), the DLPFC is

believed to exert top-down control over prepotent responses, biasing behaviour towards attending higher-level goals (Miller & Cohen, 2001). This suggests that the essence of behavioural flexibility may lie in an optimal functional connectivity between the lateral prefrontal cortex and caudate nucleus, rather than in isolated activations of these brain regions.

In contrast to clinical studies that focus on deficiencies, fewer studies have discussed superior behavioural flexibility. A noted exception is the work of Krugel et al. (2009) who investigate how genetic variation in dopaminergic modulation of the brain causes individual differences in behavioural flexibility in a healthy sample. They show that individuals with the ancestral Val/Val polymorphism of the COMT gene have higher and more flexible learning rates, leading to better performance in a learning task, compared to individuals with a Met/Met-polymorphism. Based on connectivity analysis, they furthermore suggest that optimal behavioural flexibility is accomplished through down-regulation of the ventral striatum by the prefrontal cortex. In the current study, we further investigate whether activation of this frontostriatal network coincides with self-reported self-regulation capacities and actual performance on a task challenging behavioural flexibility.

We use a probabilistic reversal learning (PRL) task to challenge behavioural flexibility in a healthy population of young males. In this task, reward contingencies between action and outcome are learned based on feedback. After a certain number of trials, the reward contingencies change without notification. The participants have to capture this change. To avoid that participants learn that reward contingencies have reversed after one episode of negative feedback, probabilistic reinforcement is introduced: in most cases a correct answer leads to reinforcement but in some cases, a correct answer is followed by negative feedback. This probabilistic nature forces the participant to integrate feedback over a series of trials and resist the urge to switch behaviour too quickly. This task has previously been

used in similar fMRI studies as a measure for behavioural flexibility (Cools, Clark, Owen, & Robbins, 2002; Cools, Sheridan, Jacobs, & D'Esposito, 2007; Ersche et al., 2011) and is proven to reliably evoke activation in the frontostriatal network during fMRI (Freyer et al., 2009). We briefly elaborate on this research in order to develop subsequent hypotheses extending previous results.

1.1. The neural correlates of reversal learning

Previous research on the neural correlates of behavioural flexibility that relied primarily on PRL paradigms found significant roles for the DLPFC, ventrolateral prefrontal cortex (VLPFC), OFC, anterior cingulate cortex (ACC) and nucleus accumbens (Nacc) (Budhani, Marsh, Pine, & Blair, 2007; Cools et al., 2002; Cools, Lewis, Clark, Barker, & Robbins, 2007; Freyer et al., 2009; Hampton, Bossaerts, & O'Doherty, 2006; Krugel et al., 2009). Each of these brain regions is thought to have a specific role in reversal learning. The Nacc, VLPFC and OFC are involved in reward processing and in learning stimulus-response associations. When the OFC fails, problems arise in monitoring the changes in the reward value of stimuli and, as a consequence, this information can no longer be used for guiding behaviour (Bachevalier & Loveland, 2006; Dias, Robbins, & Roberts, 1997). Together with the Nacc in the ventral striatum, the OFC forms part of a circuit that is responsible for encoding the value of rewards (Knutson, Fong, Bennett, Adams, & Homme, 2003; McClure, Laibson, Loewenstein, & Cohen, 2004; O'Doherty et al., 2004; Padoa-Schioppa & Assad, 2006) and for registering prediction errors when rewards are not as expected (Schultz, Dayan, & Montague, 1997). A positive prediction error signaled in the Nacc strengthens the association between stimulus and response.

The DLPFC is typically recruited in tasks that require exerting cognitive control (Balleine & Dickinson, 1998), working memory (Leh, Ptito, Chakravarty, & Strafella, 2007) and attentional control (Hampshire,

Chaudhry, Owen, & Roberts, 2012), all of which may be needed to monitor trial-to-trial changes and arrive at the new stimulus-response association.

Finally, the ACC is known to be involved in value estimation, detection of errors and prediction errors as well as conflict monitoring (Silvetti, Seurinck, & Verguts, 2011). Specifically relevant to reversal learning, activation of the ACC may be representative of the cost associated with a decision, as it consistently responds to pain, error monitoring, response conflict, and even mental effort (Botvinick, Cohen, & Carter, 2004; Carter et al., 1998; Cohen et al., 2007). Negative feedback following errors has specifically been shown to elicit activity in the posterior rostral cingulate zone (RCZp), the more dorsal part of the ACC (Hampshire et al., 2012). Second, the anterior rostral cingulate zone (RCZa), a more ventral part of the ACC, is likely involved in representing and updating the relationship between action and outcome (Carter et al., 1998; Hampshire et al., 2012). This latter function in updating behaviour of the ACC has also been attributed to the caudate nucleus (Grahn, Parkinson, & Owen, 2008).

Surprisingly, the caudate nucleus, located in the dorsal striatum and part of the basal ganglia, has not been previously implicated in accomplishing successful switching in PRL. Nevertheless, several functions of the caudate nucleus suggest that it might be: it is involved in cognitive processes requiring agency (Baumgartner, Heinrichs, Vonlanthen, Fischbacher, & Fehr, 2008) and in updating behaviour (O'Doherty et al., 2004), based on reinforcement learning (Tricomi & Fiez, 2008) and reversal learning (Clatworthy et al., 2009; Tricomi, Delgado, McCandliss, McClelland, & Fiez, 2006). By comparing a Pavlovian and instrumental learning task, O'Doherty et al. (2004) have described how the ventral and dorsal striatum (the latter including the caudate nucleus) are dissociated. The caudate nucleus becomes only activated during instrumental learning (associating a stimulus with a response), and not during a Pavlovian task (associating stimulus with a reward). Thus, an active caudate nucleus would indicate that actions with a

better (or worse) outcome will tend to be repeated (or attenuated) in next trials. A possible reason for the absence of caudate nucleus activation in PRL tasks is that the pattern of activation might be dynamic, increasing with perseverating errors, and that it would therefore not show up in simple contrasts that isolate a moment in decision-making (see also King-Casas et al., 2005).

1.2. Hypotheses

The goal of this study is to extend previous research on PRL and to further examine individual variation in neural activation in a healthy population. Our specific goal is three-fold. First, we replicate previous findings (Cools et al., 2002; Cools, Lewis, et al., 2007) reporting DLPFC, VLPFC, OFC, ACC and Nacc during trials in which participants decide to switch behaviour (following negative feedback), compared to correctly executed trials (followed by positive feedback). We fine-tune this finding by investigating whether a more stringent contrast between errors that lead to switching, and previously made errors that do not lead to switching, activates similar brain regions. In doing so, we isolate the moment of insight that leads to a switch in behaviour, and we eliminate possible noise related to the valence of feedback (negative after the last error that leads to switching, positive after correct trials).

Second, we specifically test for the involvement of the left and right head of the caudate nucleus during PRL. Given the evidence that the caudate associates a stimulus with the outcome and hence may lead to perseveration in instrumental learning, we investigate the possibility that the caudate activity increases with the number of perseveration errors, i.e., with the number of consecutive feedback errors prior to switching. Due to the probabilistic nature of the task, a single instance of negative feedback after a series of corrects answers may not be a sufficient indicator to switch. Only when an individual perseverates in making errors (and negative feedback accumulates in subsequent trial(s)), caudate activity is expected to increase. The association between negative feedback and behaviour would then

be strengthened until the moment of switching. Therefore we will test if feedback errors prior to switching parametrically modulate the BOLD (Blood Oxygenation Level Dependent) signal in the caudate nucleus.

Third, we investigate if brain activity (BOLD signal) during PRL varies as a function of individual differences in self-regulation in a normal (non-clinical) population of young adults. To assess normal individual variation in self-regulation, we rely on two different measures. First, we preselect high and low self-regulating participants for this experiment based on their self-reported scores on the Effortful Control scale of the Adult Temperamental Questionnaire (ATQ). This measure is not related to performance on the task. Next we also use a behavioural measure of self-regulation based on participants' actual performance during the PRL task. We hypothesize that self-reported self-regulatory abilities will be associated with improved performance on the PRL task, and that this is furthermore associated with enhanced cognitive control, i.e., increased activation in the PFC and dorsal ACC (Miller & Cohen, 2001).

2. Method

2.1. Population

40 healthy right handed men (mean age 19.83, SD = 1.45; range = 18-24) were recruited to voluntary participate in the study, which was introduced as an investigation of the brain regions involved in decision-making. None of the participants had a history of neurological or psychiatric illness, were on any medication, or suffered from claustrophobia. All candidates gave written informed consent and received a monetary compensation ranging between € 25 and € 40 for their participation which included a show up fee (€ 15) and a sum that depended on the choices they made during the task. This

experiment was part of a larger study assessing various components of self-regulation with multiple tasks[1].

2.2. General procedures

Candidates were invited by email to participate in the study and were asked to complete an fMRI screening questionnaire and a personality questionnaire. We quantified individual differences in self-regulation a priori with the Effortful Control scale of the Adult Temperament Questionnaire (ATQ) (Rothbart, Ahadi, & Evans, 2000) specifically developed to assess stable, temperamental differences in healthy, young adults (Evans & Rothbart, 2007). The questionnaire comprises three subscales "inhibitory control", "attentional control" and "activational control." A total of 211 university students filled in the Effortful control scale of the ATQ. Only those individuals who scored above (high self-regulators (HSR), n = 25) or below (low self-regulators (LSR), n = 15) the 50th percentile for all three subscales of the Effortful Control scale were invited to participate in a single scan session at the Antwerp University Hospital.

Upon arrival on the day of the experiment, the participant signed an informed consent form and next received instructions for the PRL task To get fully acquainted with the task, 20 practice trials were conducted on a laptop prior to the actual experiment. Next, the participant was positioned in the scanner. Anatomical data was acquired during the first 10 minutes, followed by three functional runs, each lasting 9 minutes. In between functional runs there was a short break (< 2 minutes) but the participant did not leave the scanner. Stimuli were presented using E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA). An fMRI compatible screen was positioned at the end of the scanner to

[1] All participants completed two separate tasks under the scanner. First they engaged in a time discounting task which assessed their ability to delay gratification, see Waegeman, Declerck, Boone, Van Hecke, and Parizel (2014). After completing the time discounting task, participants left the scanner for approximately 15 minutes during which they received instructions for the PRL task. Because of the reported differences in reward valuation during incentive delay tasks between the sexes (Spreckelmeyer et al., 2009), we restricted the sample for the entire study to males.

depict the projected stimuli to the participant, lying in the scanner with a mirror installed on the head coil and holding a response button box in their right hand to express their choice. Effort was made to restrict motion during scanning (e.g., foam cushions were inserted between the head and the head coil) so that less than 2 mm translational or 2°rotational head motion was recorded within a run. At the end of the experiment the participant was remunerated for his participation. The total amount was dependent on the choices made during the task (see below). Participants earned between € 2.75 and € 20.8[2]. Above procedures were approved by the Commission of Medical Ethics at the University of Antwerp.

2.3. Behavioural task

The task is adapted from the PRL task used by Cools et al. (2002). A single trial of the task is illustrated in Figure 1. Two stimuli (letters "A" or "B") were presented simultaneously on the left and right side of the center of a screen. Participants were told that they had to make a choice between A or B upon which they received feedback ("correct" or "incorrect"). A correct choice yielded a reward of € 0.25 while incorrect meant the participant lost an amount of € 0.25. The participants were explicitly told to earn as much money as possible. The stimuli (A and B) were presented for a maximum of 2000 ms, the choice was underlined for 500 ms. This was followed again by the feedback signal for 500 ms, after which a fixation cross was presented for a variable duration between 500 ms and 1500 ms. When participants did not respond within 2000 ms, a message reading "too late" appeared on the screen for 1000 ms, followed by a fixation cross of variable duration. Participants played three runs of 9 minutes during which they had to continuously make choices.

[2] As explained in the informed consent, participants were only paid one third of the amount earned throughout the experiment, which ranged from € 8.25 to € 62.5.

The task was programmed in such a way that at the start of the experiment, option “A” led to a reward in 80% of the cases. In 20% of the cases, option “A” was followed by negative feedback. This contingency between stimulus and reward changed without informing the participant so that suddenly “B” led to a reward in 80% of the cases. This reversing in contingencies behind A and B is called a reversal stage shift. A reversal stage is a series of trials in which the stimulus-reward contingencies do not change. The number of trials after which the contingency does change was dependent on the number of consecutive correct trials the participant achieved. To avoid that participants would easily figure out this number, E-prime randomly picked a number between 10 and 15 as the number of correct trials the participant had to achieve before a shift in reversal stages occurred. Each reversal stage thus has a variable length. The total number of reversal stages achieved during the task, reflected how rapidly the participant gained insight into the necessity to adjust behaviour. Thus, a participant who accomplished many reversal stages quickly picked up the reversal in contingency and was better able to stop responding to the previously learned rule. A low number of reversal stages meant that it took the participant longer to discover that the rule had changed and/or took the participant longer to stop responding to a previous learned rule. Therefore the total number of reversal stages is a meaningful behavioural measure of flexibility. Individuals who achieve many reversal stages are switching optimally, while a small number of reversal stages can be indicative of switching either too soon or too late, as both eventually lead to fewer correct answers.

Participants made their choice by pressing one of two buttons on the button box which they held in their right hand during the experiment. The left button represented the left choice on the screen, the right button represented the right choice on the screen. Presentation of the choices was randomized so A or B changed position on the screen regularly.

Following the methods used by Cools et al. (2002; 2007) we define the following different types of trials: a correct trial, a probabilistic error (PE), a reversal error (RE) and a final reversal error (FRE). A correct trial was defined as giving the correct answer followed by positive feedback. A trial labeled as a PE meant negative feedback was given after a correct answer. Trials in which positive feedback was given after a wrong answer were also modeled as a PE, but these trials were of no interest with respect to the aim of this study and therefore excluded in further analyses. The inclusion of PEs makes the task more difficult (Waltz & Gold, 2007). Without PEs, the contingency reversals would be too easy to detect. Participants were not told about the occurrence PEs. In this respect, the instructions participants received with respect to rule reversals differ slightly from those that were given in the experiments by Cools et al. (2002; 2007) where subjects were instructed "to only start choosing the other pattern when they were sure that the rule had changed", (2002, p. 4564; 2007, p. 182). In this study, participants were merely told to "earn as much money as possible."

An RE trial is defined as a trial in which the participant gave an erroneous answer which was the same as the answer that was given in the previous trial. Feedback after a RE trial is always negative. Finally, for a trial to be labeled as an FRE, three conditions needed to be fulfilled: (1) the incorrect answer of the previous trial that yielded negative feedback was repeated, (2) the trial was followed by a behavioural switch in the next trial, (3) the subsequent correct answer was sustained for at least one more consecutive trial after the switch. We defined an FRE in such a way to make sure that a random answer would not be mistaken for an insightful behavioural switch. As such, an FRE is the last error in a series of perseverating errors and represents the moment after which participants realized their behaviour was no longer adaptive and had to be changed in order to continue to earn money.

2.4. Image Acquisition

Data were collected with a 3 Tesla Siemens Trio MRI scanner system (Siemens Medical Systems, Erlangen, Germany) at the Antwerp University Hospital, using a 32-channel radiofrequency head coil. First, 176 high resolution anatomical images were acquired using a T1-weighted 3D MPRAGE sequence (TR = 1910 ms, TE = 3.37 ms, flip angle 15°, image matrix = 192 x 256, FOV = 256 mm, voxel size = 1x1x1 mm). In each run, whole brain functional images were collected using a T2*-weighted EPI sequence, sensitive to BOLD contrast (TR = 2000 ms, TE = 50 ms, flip angle 90°, image matrix = 64 x 64, FOV = 192 mm, voxel size = 3x3x3 mm, 31 sagittal slices, slice thickness = 3 mm, interslice gap = 0 mm, interleaved scanning mode). To ensure accuracy of event timing, E-Prime 2.0 was programmed to start when triggered by the EPI pulse so the start of the stimuli presentation was synchronized with the registration of the first volume. The first three volumes were discarded to allow for T1 equilibrium effects.

2.5. Data analysis

Image analysis was conducted in SPM8 (The Wellcome Trust Centre for NeuroImaging, London). Preprocessing included slice time correction, realignment, coregistration of anatomical and functional images, normalization to Montreal Neurologic Institute (MNI) coordinates and spatial smoothing with a FWMH Gaussian kernel of 8 mm. The dependent variable in our analyses is the BOLD response.

To identify the brain regions that are activated at the moment of switching (see the first hypothesis), a general linear model (GLM) was created for each participant, which included three regressors of interest, referring to the following trials: a correct trial, RE and FRE. Variance in the data due to other types of error (i.e., PE, late trials, and motion) was controlled for by including them as factors in the GLM. Because several brain regions are still maturing during young adulthood (the sample population) we also included age as a covariate in the GLM. Regressors of interest were created by convolving a box

car function representing the onset and duration of a trial with a canonical hemodynamic response function. The event started when the stimuli were displayed and ended after the feedback was given. Variation in trial duration was dependent on an individual's reaction time. Two whole brain contrasts were calculated at a threshold of $p < 0.05$ (Family Wise Error (FWE) corrected): (1) FRE > correct trials (replicating the work of Cools et al., 2007) and (2) the more stringent contrast, FRE > RE, which isolates the awareness that the reward contingencies have reversed.

To specifically test for the neural correlates of perseveration (second hypothesis), we computed a second GLM similar to the previous one but including parametric modulation of the RE trials. This parameter refers to the position in a series of consecutive REs leading to the FRE. So the more perseverative errors or consecutive REs are made, the higher the parameter becomes. This parametric modulation of reversal errors is contrasted with baseline. First we estimate this contrast for the entire brain at $p(FWE) < 0.05$. Next we perform a region of interest (ROI) analysis around the caudate heads as predefined anatomical structures in the WFU PickAtlas toolbox of the Wake Forest University School of Medicine (Maldjian, Laurienti, Kraft, & Burdette, 2003). Peak value coordinates were used for localization using the Automated Anatomical Labeling (AAL) toolbox for SPM.

To test for individual differences in the neural basis of switching (hypothesis 3), we repeat both the FRE > RE contrast and the parametric modulation of RE > baseline contrast, but each time we add a measure of self-regulation as a covariate. First we added the ATQ scores (self-report measure) as a covariate in the regression. Next we repeated the analysis with the number of reversal stages that each participant achieved during the task.

3. Results

3.1. Behavioural data

Participants (n = 40) completed anywhere between 25 and 38 reversal stages (mean = 32.22, SD = 2.9), and made between 36 and 61 FREs (mean = 46.33, SD = 6.1). An advantage of relying on the total number of reversal stages as a measure of individual differences is that this number reflects the efficiency of the strategy of a participant: the number is high when a participant is not fooled by a PE but instead persists on providing the correct answer, reducing the length of a reversal stage. A disadvantage of relying on the total number of reversal stages as a measure of individual differences is that good performers may possibly also react more quickly and consequently may have more trials and more spreading of the error. Table 1 compares good performers with bad performers , showing for each group the mean reaction times, mean total number of correct trials, mean total number of RE and mean total number of FRE. We additionally perform independent sample t-tests on these data (see Table 1). Mean reaction times differs significantly ($t = 3.068$, $p = 0.006$). Good performers work faster. Note that neither the total number of RE, nor the total number of FRE, is significantly different between the groups, suggesting that it is not the frequency of switching, but switching at the correct time that is crucial for success.

Table 2 shows that, unlike expected, the total number of reversal stages did not correlate with the effortful control score (Pearson's $r = -0.078$, $p = 0.631$). As expected, the number of reversal stages correlates significantly with the total number of correct trials (Pearson's $r = 0.930$, $p < 0.01$), the total number of trials (Pearson's $r = 0.564$, $p < 0.01$) and total earnings in the probabilistic reversal learning task (Pearson's $r = 0.930$, $p < 0.01$), see Table 2. As shown by Krugel et al. (2009), good performers (those individuals with a high number of reversal stages) also have a higher and more flexible learning rate. Therefore, in addition to working faster, they may also be less side-tracked by probabilistic errors

(which may occur within a series of RE leading up to a reversal stage) and switch more effectively. The choices they make lead to a greater number of correct trials, spread over a greater number of shorter reversal stages. Neither the total number of RE, nor the mean number of RE preceding FRE correlates significantly with the number of reversal stages (Table 2), indicating that perseveration by itself is not related to performance. The maximum range of RE before FRE ranges from 5 to 14 for HSR and from 4 to 15 for LSR.

3.2. Imaging data

3.2.1. Whole brain analysis

Brain regions that survive p(FWE) < 0.05 correction for the contrast FRE > correct trials and FRE > RE are reported in Table 3. The results of the first contrast (left side of Table 3 and Figure 2) replicate previous research (Cools et al., 2007) by finding activation in bilateral DLPFC, inferior frontal gyrus, frontal eye fields, bilateral VLPFC, parietal and temporal cortex and right dorsal ACC. Unlike Cools et al. (2007), we do not find right OFC or Nacc[3]. Eight of eleven clusters survive the more stringent contrast FRE > RE (see bold printed regions in Table 3 and right side of Figure 3. In this contrast, the peak activation of the dorsal ACC has shifted to a more rostral position. The DLPFC is only observed on the right side of the brain. As in previous research, no caudate activation is identified in these whole brain contrasts, not even when we relax the statistical threshold to p < 0.001, uncorrected, or when we perform a separate ROI analysis, using the WFU atlas predefined shapes of the caudate heads.

3.2.2. Parametric modulation

Table 4 reports the results of the parametric modulation of RE contrasted against baseline. Only the bilateral dorsal ACC survives the p(FWE) < 0.05 statistical threshold. Further investigation with an ROI

[3] When relaxing the statistical threshold to p < 0.001, uncorrected, we still do not observe OFC or Nacc. This is probably due to the susceptibility artefact in the functional images at the location of these brain regions.

analysis (using WFU atlas' anatomically predefined shapes of the caudate heads) corroborates that the right caudate head becomes more activated during perseveration (i.e., with increasing reversal errors), see Table 4 and Figure 4.

3.2.3. High self-regulators versus low self-regulators

For the FRE > RE contrast, two brain regions correlate positively with the number of reversal stages at p(FWE) < 0.05: the right rACC (with a cluster size of 372 functional voxels at x = 3, y = 23, z = 37) and the right VLPFC (with a cluster size of 89, x = 33, y = 23, z = -14). This indicates that, when they are aware of a reversal in reward contingencies, these two regions are more activated in those individuals that perform better on the task. Figure 5 illustrates that individual differences in performance on the current task are also linearly related to activation in the ACC and the VLPFC. The number of reversal stages did not correlate with any brain regions (at p(FWE) < 0.05) in the parametric modulation of reversal errors, suggesting that the brain regions involved in perseveration in PRL do not vary with performance ability. With respect to the self-report measure, no brain regions in either contrast (FRE > RE and parametric modulation of RE > baseline) were found to correlate positively or negatively with self-regulatory ability.

4. Discussion

The results of this study fit in well with current knowledge on the neural underpinnings of self-regulation. Before elaborating on the specific contributions, we draw attention to two technical constraints in the design of the experiment which have posed limitations on the interpretability of the data and, subsequently, on the conclusions that can be drawn. First, the absence of multiple jitter periods during a given experimental trial precludes the differentiation of the isolated events that are progressively unfolding. Thus, stimulus evaluation, choice, prediction, or prediction error during feedback cannot be disentangled and therefore the reported BOLD activation can only be attributed to

overall decision-making during a particular trial. The implication of this limitation is that we could not further analyze the neural basis of individual differences in learning rates with a solid computational model. As such, this analysis is now limited to correlating an overall performance measure (the total number of reversal stages reached by an individual) with brain activity extending over an entire trial.

The second technical constraint of the design is the short jitter time, varying from 500 to 1500 ms. This time corresponds to the variable intertrial interval (during which a fixation cross appeared on the screen) which allows to desynchronize from the repetition time and to sample sufficiently across the hemodynamic response function. A longer intertrial interval would have been necessary in order to be able to reliably infer effective connectivity between the involved brain regions. Hence, given the current data, we can also not derive any conclusions regarding the directionality of the activity.

Finally, we note again that the current study was limited to young male participants. While this has several advantages, such as eliminating additional noise in the data due to gender differences, the current conclusions should not be extended to the entire population at large. Brain anatomy and chemistry are known to vary with age and gender and cause differences in processing tasks such as the PRL task (Robinson, Standing, DeVito, Cools, & Sahakian, 2010).

Despite these limitations, the current study contributes to the extant literature in three meaningful ways. First, we corroborate previous results of fMRI studies by showing that successful switching in a PRL task involves activation of a dominantly right lateralized cognitive control network that includes the dorsal ACC, DLPFC, VLPFC, frontal eye fields, and parietal cortex (Cools et al., 2002; Freyer et al., 2009). In addition, by contrasting FRE with RE instead of correct answers we succeeded to isolate the awareness of a reward contingency reversal in a more stringent manner, eliminating all possible confounds related to the valence of feedback. Interestingly, the peak activation in the ACC shifts from a dorsal position (FRE > correct) to a more rostral position (FRE > RE). However, because the contrast FRE

> correct also includes the peak coordinate of FRE > RE, we cannot be certain that the shift represents a functionally or architecturally different brain area. The activation identified in the rACC might also represent a more ventrally extended part of the dorsal ACC. The dACC is a very heterogeneous region; with its known connections to the lateral prefrontal cortex, parietal cortex, and frontal eye fields (see also Table 3), it is ideally located to serve as an information integration center to directly influence attention allocation, motor preparation, and motor responses (Bush, Luu, & Posner, 2000; Bush et al., 2002). However, both dACC and rACC are known to become activated during error commission, yet with different dynamics: the dACC appears to become involved immediately after a response error and contributes to the early updating of stimulus-response mapping, while rACC activity rises more gradually and peaks at a later time during the evaluative process reflecting the affective appraisal of the error. Performance optimization would then depend on contributions of both the rACC and the dACC (Polli et al., 2005). This is consistent with what we observe in the current study, with the rACC only becoming more active during an FRE, and less so during other REs.

Dorsolateral prefrontal areas are specifically involved in cognitive set shifting (Freyer et al., 2009), whereas the right VLPFC reflects behavioural inhibition (Cools et al., 2002). Moreover, the right VLPFC has been shown to play a role in the down-regulation of negative emotional responses (Tabibnia, Satpute, & Lieberman, 2008). In the PRL task, the accumulation of negative affect as a result of continuous negative feedback during a series of reversal errors could possibly also be down-regulated as indicated by the increased activation of the right VLPFC. This would mean that the participant experiences increased resilience to override the default response (which is no longer correct due to a contingency reversal). Similarly, the bilateral DLPFC, rACC, bilateral insula and left inferior parietal lobe (angular gyrus) are 6 of the 15 cortical areas that were also reported to have increased activation in a study by Sanfey et al. (2003) when participants under the scanner were treated unfairly in an economic

game. A subsequent study showed that VLPFC activation was necessary to respond rationally to unfairness and to control the emotional impulse to punish the wrongdoer (Tabibnia et al., 2008).

A region that is anatomically connected and functionally related to the insular regions is the superior temporal gyrus (Paulus, Feinstein, Leland, & Simmons, 2005). Together with insular regions, the part of the superior temporal gyrus at the border with the angular gyrus is involved in integrating information that occurs over time. A study using a Rock, Paper Scissors game reports both regions as important for integrating the previous response and previous outcome to subsequent decision making. This contextual information provided by the superior temporal gyrus is integrated in other brain regions such as the insula and is critical to decide whether or not ongoing actions should continue (Paulus et al., 2005). Activation of the superior temporal gyrus thus adds to the action – outcome dependent response strategy, critical in reversal learning.

The involvement of the precuneus and left inferior parietal lobe (angular gyrus) in implementing the switch (contrast FRE > RE) is consistent with previous reports identifying those regions as being important in first-person perspective taking and experiencing agency (Ruby & Decety, 2001). The fact that in the current study these regions become more activated immediately preceding a behavioural switch, suggests that the participant is self-reflecting on his/her own (deficient) behaviour and taking charge of fixing it.

Taken together, the regions identified in Table 3 show a good fit with the different attentional networks that have been described for alerting, orienting, and executive control (Fan, McCandliss, Fossella, Flombaum, & Posner, 2005). All three networks might be orchestrated during performance in the PRL in the following way: First, the alerting network (consisting primarily of right frontal, (inferior) parietal and superior temporal regions) becomes activated following negative feedback and alerts that “I have to do something.” Next the orienting network (centered on the superior parietal areas, including the

precuneus) responds to the alert by reflecting on “what are my options.” This is followed by engaging the executive control network (comprising the ACC and the LPFC) which signals “let’s change this!” Overall, the involvement of the latter network strengthens the idea that executive functions to achieve successful behavioural flexibility in a PRL task rest on strong involvement of the ACC along with the concerted activity in the lateral PFC.

The second contribution is the finding that the dACC appears to be the most active region during perseveration, and the only cluster large enough to survive statistical thresholding (see Table 4), which is consistent with its role in monitoring and updating action – outcome contingency (Carter et al., 1998; Hampshire et al., 2012). Alexander and Brown (2011) presented a unifying model of ACC functioning, suggesting that ACC is concerned with learning and predicting the likely outcomes of actions. More specifically, negative surprise signals in ACC (accumulated during perseveration) would hold information about context-specific predictions and evaluations (Alexander & Brown, 2011). Because the heads of the caudate nucleus are also known to be activated during contingency learning and updating behaviour (Grahn et al., 2008), we assessed their involvement with an ROI analysis and found evidence that the right caudate nucleus also becomes more activated with increasing reversal errors. The fact that caudate head activity can only be identified when including RE as a parametric modulator in the GLM confirms that learning in a PRL task is a dynamic process that extends over several trials. We speculate that, by strengthening the association between negative feedback and response, the caudate nucleus contributes to the realization that behaviour is no longer adaptive. Along those lines, an fMRI study by Baumgartner et al. (2008) reports that individuals with reduced caudate activity were less able to adapt their behaviour after receiving consecutive negative feedback from their partners with whom they were playing an interactive game. This realization that behaviour is no longer adaptive is key to the behavioural switch that follows, but cannot, however, be entirely dependent on activation of the

caudate head since no individual differences are noted in caudate activation. The actual signal to change may reside elsewhere.

Third, with respect to individual differences in the neural correlates of behavioural switching, the rACC and the right VLPFC appear to be the most important regions that underlie successful behavioural flexibility, indicating that cognitive control achieved by the joint activity of the lateral PFC and the error-monitoring faculties of the ACC (Miller & Cohen, 2001) is recruited more intensely by good performers in the PRL task. Moreover, Krugel et al. (2009) find that activity in the PFC is functionally connected to and down-regulates activity in the ventral striatum during a behavioural flexibility task. This corroborates the extensive literature reporting individual variation in executive functions (Hofmann, Schmeichel, & Baddeley, 2012) of which behavioural flexibility is one of them. In contrast to our hypothesis stated in the introduction, the caudate nucleus does not correlate with individual differences in task performance in the whole brain analysis. Interestingly, however, frontostriatal connections have been reported to be essential in successful PRL by Krugel et al. (2009). By investigating the influence of a genetic polymorphism in the enzyme catechol-O-methyltransferase (COMT) on the neural correlates of behavioural flexibility they convincingly showed that individual differences in flexibility in reinforcement learning are related to differences in response to prediction errors (the discrepancy between reward prediction and reward occurrence) in the ventral striatum. They find that Val homozygotes have a greater and a more differentiated ventral striatal response to prediction errors, resulting in higher and more flexible learning rates during PRL. Val homozygotes also show greater learning rate-dependent changes in the effective connectivity between ventral striatum and PFC, as a result of a downstream effect of prefrontal dopamine levels (Krugel et al., 2009). Their effective connectivity analysis shows that down-regulation of the striatal response by the PFC underlies individual differences in behavioural flexibility. Consistent with their evidence that PFC activity underlies individual differences, we find in the

current study that the BOLD response in the rACC and the VLPFC correlates positively with increasing behavioural flexibility.

A final and major additional finding which was not expected is the absence of a consistent relation between brain and behaviour on the one hand, and self-reported self-regulation on the other hand. Although we specifically opted for a questionnaire that is designed to characterize effortful control in young healthy populations, it is possible that healthy individuals are hard to differentiate on the basis of such a questionnaire or that adolescents may not have reached stable levels of trait self-regulation that can be accurately assessed with questionnaires. In addition, temperamental self-regulatory capacity (trait self-regulation) measured by self-report probably is not necessarily related in a one to one way to state dependent self-regulation measured during the task. As such, reporting to have the capacity to self-regulate and to actually do so during a laboratory task a few weeks later may be confounded by differences in mood or other state dependent fluctuations in preferences.

In conclusion, adding to the existent literature, the results of the current study corroborate and provide additional insights into the neural correlates of performance during PRL. Prior to a behavioural switch, dACC and right caudate become increasingly activated with accumulating errors. At the moment of switching, rACC, VLPFC and DLPFC become more activated, with the former two correlating positively with performance on the task. As an avenue for future research, dynamic causal modeling analysis could reveal whether or not the strength of a frontostriatal pathway connecting the caudate nucleus with the PFC accounts for variation in behavioural flexibility and whether this is the case for both healthy and clinical populations.

What's your choice?	Your choice :	Feedback	Fixation :
A or B	A B	Incorrect!	+
Duration : max 2 s	500 ms	500 ms	variable

Figure 1 illustrates an example of a single trial in the PRL task.

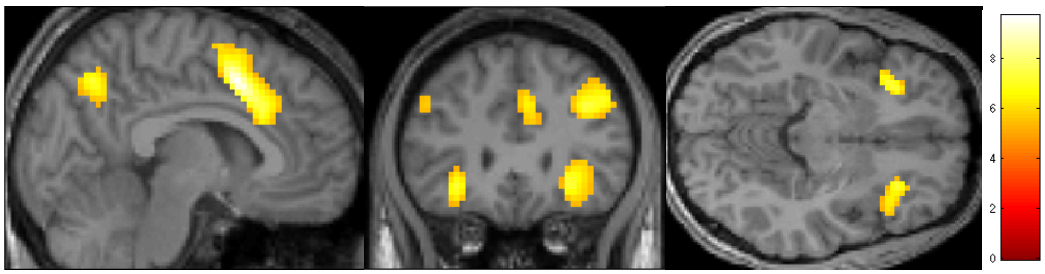

Figure 2 shows the p(FWE) < 0.05 corrected brain regions that are activated in the contrast FRE > correct. From left to right: activation of ACC (at x = 7, y = 25, z = 2), activation of bilateral insula/VLPFC and DLPFC (at x = 7, y = 25, z = 1) and finally left insula/VLPFC (at x = 7, y = 25, z = -8) from an axial viewpoint. Bar on the right denotes T values of colored regions.

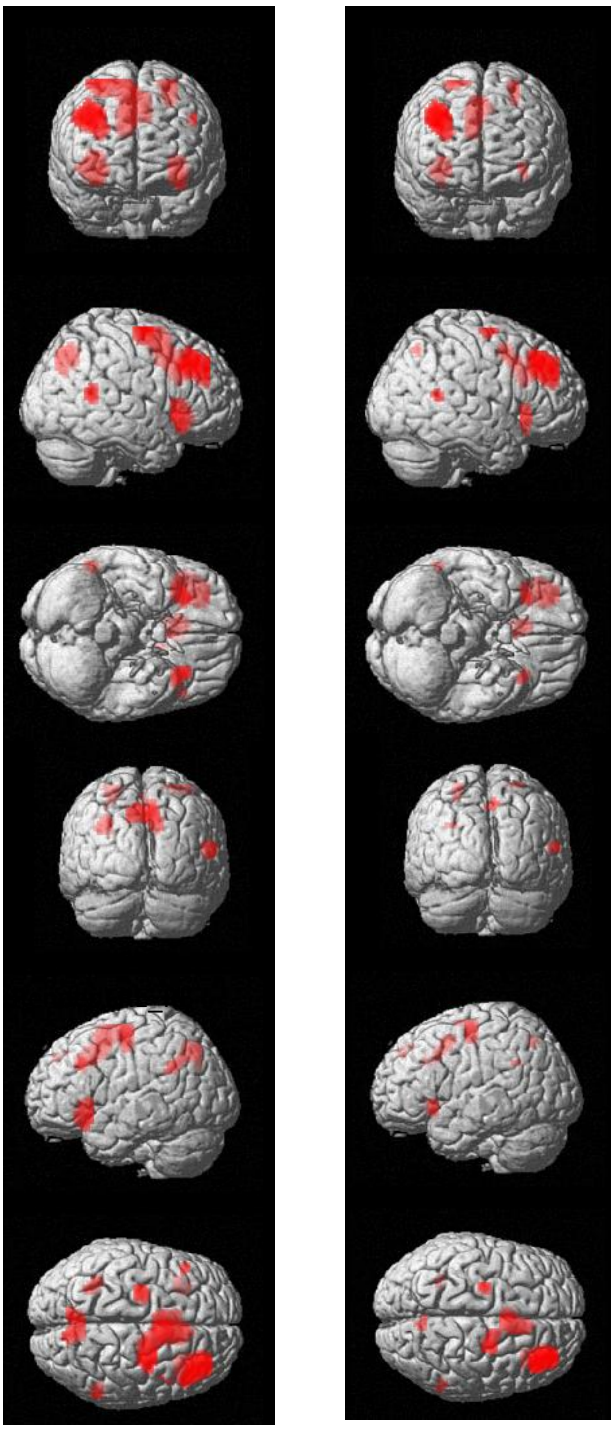

Figure 3 compares the activated regions for the contrast FRE > correct (left panel) with FRE > RE (right panel) at p(FWE) < 0.05. Both contrasts result in a similar network. Cluster sizes in the more stringent contrast (right panel) are smaller than the original contrast (left panel).

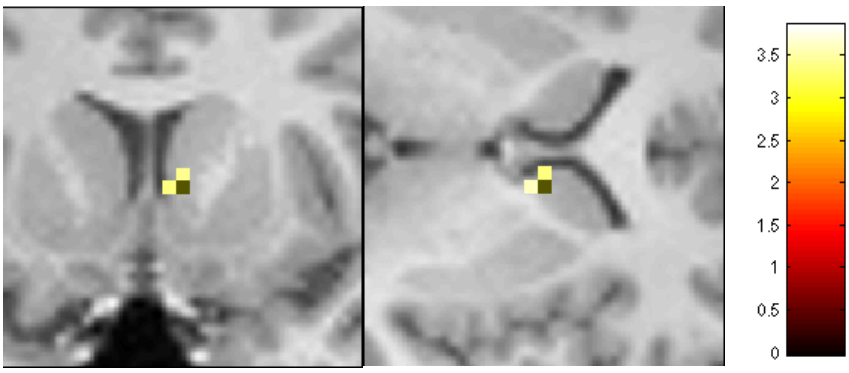

Figure 4: The contrast parametric modulation of RE > baseline shows activation of 4 voxels in right caudate after ROI analysis in both caudate nuclei at p(FWE) < 0.05, with a peak at x = 9 y = 8, z = 4. Image is zoomed 80*80mm. Bar on the right denotes T values (peak at 3.85).

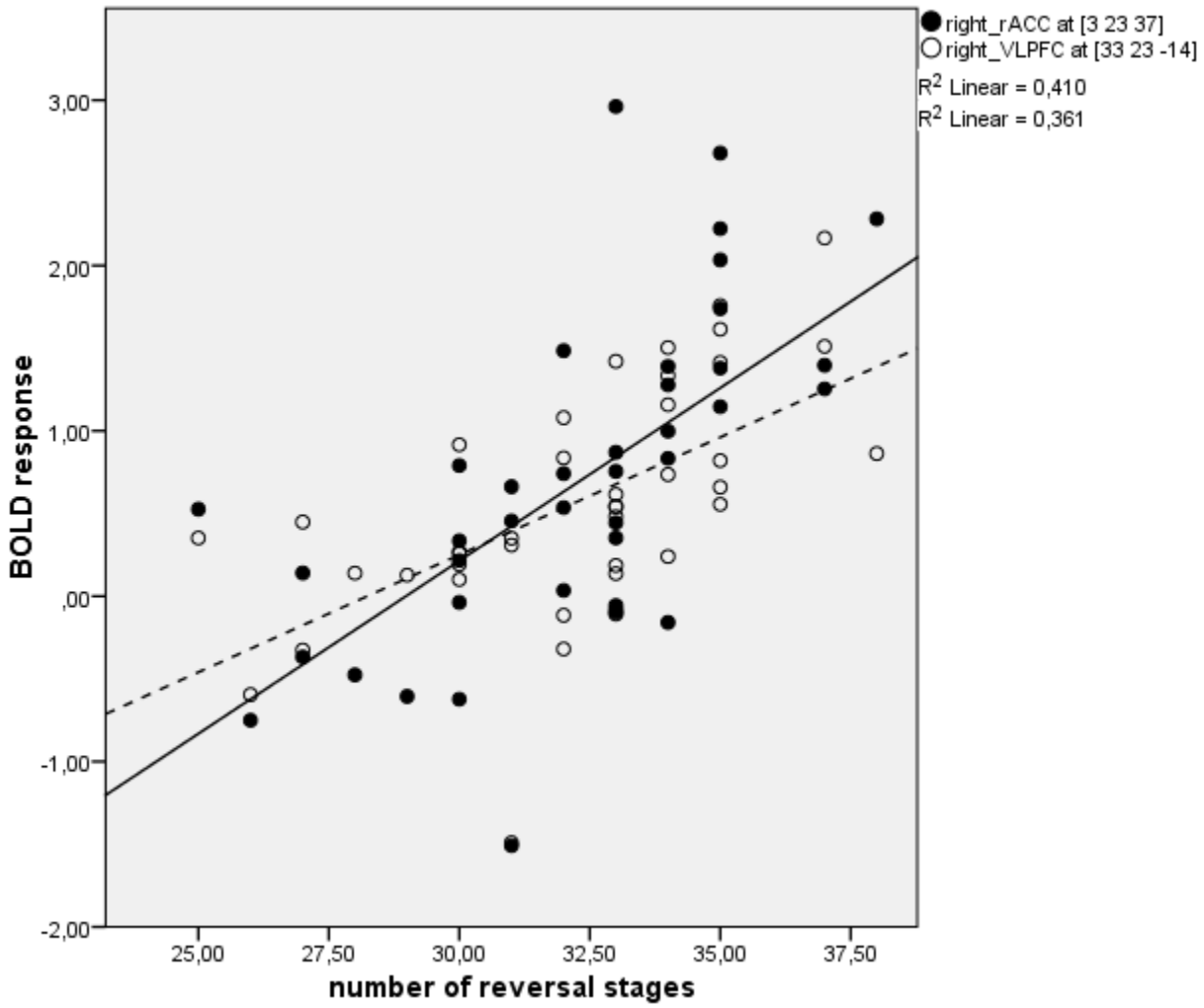

Figure 5 shows the correlation between the number of reversal stages and the BOLD response at right rACC (full dots) and right VLPFC. Mean beta values were extracted from the FRE > RE contrast with 8 mm spheres round MNI coordinates x = 3, y =23, z = 37 for rACC and x = 33, y =23, z = -14 for VLPFC. Response in both regions increases linearly with performance (rACC R^2(full line) = 0.410 and VLPFC R^2(dotted line) = 0.361).

Table 1: Independent samples t-tests between low - and high self-regulators on behavioural variables of the probabilistic reversal learning task

	LSR	SD	HSR	SD	t	Standard error
Mean reaction times	638.26	140.19	531.60	50.66	3.068**	34.76
Total reversal stages	29.61	2.173	34.36	1.465	-8.233**	0.577
Total correct	291.28	27.373	339.95	15.625	-6.704**	7.261
Total RE	139.11	20.384	132.05	24.699	0.972	7.268
Total FRE	46.89	5.969	45.86	6.334	0.523	1.962

LSR = Low Self-Regulators (n = 18), HSR = High Self-Regulators (n =22), SD = standard deviation, ** = $p < 0.01$ (two-tailed). Mean Reaction times are in ms. Negative t-values refer to higher values in HSR.

Table 2: Correlational analysis of behavioural variables of the probabilistic reversal learning task

	Total Reversal stages	Total correct	Total RE	Mean RE before FRE	Total trials	Money earned
Self-reported self-control	-0.078	-0.207	0.070	0.119	-0.076	-0.212
Total reversal stages	1	0.930**	-0.244	0.044	0.564**	0.930**
Total correct		1	-0.205	0.030	0.666**	0.946**
Total RE			1	0.809**	0.221	-0.348*
Mean RE before FRE				1	0.133	-0.001
Total trials					1	0.426**

Pearson correlation coefficients, n = 40, * = $p < 0.05$, ** = $p < 0.01$ (two-tailed)

Table 3: Whole brain contrasts at the moment of switching during probabilistic reversal learning

		Contrast FRE > correct						Contrast FRE > RE					
Region	L/R	x	y	z	size	T	p(FWE)	x	y	z	size	T	p(FWE)
Frontal													
dACC	R	6	8	46	579	9.73	0.000						
rACC	R							9	20	28	151	6.98	0.000
DLPFC	R	36	32	34	327	7.69	0.000	33	38	34	268	8.05	0.000
DLPFC	R	48	11	34	1	5.15	0.046						
DLPFC	L	-45	26	34	11	5.68	0.014						
Inferior frontal gyrus	R	45	5	37	1	5.26	0.035						
Frontal sup BA6	R							18	-7	67	44	6.84	0.000
Frontal eye fields	L	-24	-13	58	75	7.19	0.000	-24	-13	52	34	6.27	0.002
Insula/VLPFC	L	-33	20	-5	102	7.19	0.000	-33	20	-5	17	5.52	0.017
Insula/VLPFC	R	27	26	-8	162	7.11	0.000	30	23	-5	50	5.86	0.007
Parietal													
Precuneus	R	6	-67	46	159	7.08	0.000	6	-64	49	9	5.37	0.025
Angular gyrus	L	-30	-52	31	32	6.22	0.003	-30	-52	31	3	5.30	0.030
Temporal													
Temporal superior	R	57	-49	13	37	6.54	0.001	57	-49	10	10	5.62	0.013

Note: L = left, R = right side of the brain. Bold printed regions are shared regions over both contrasts. Coordinates are in MNI space. Cluster size represents the number of statistically significant functional voxels. All brain regions survive Family Wise Error correction $p < 0.05$.

Table 4: Parametric modulation of the number of reversal errors

		Contrast parametric modulation of RE > baseline					
Region	L/R	x	y	z	size	T	p(FWE)
Whole brain							
dACC	R	9	5	49	383	5.53	0.02
dACC	L	-6	5	49	"	5.52	0.02
ROI							
Caudate head	R	9	8	4	4	3.85	0.01

Note: L = left, R = right side of the brain. Coordinates are in MNI space. Cluster size represents the number of statistically significant functional voxels.

References

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience, 14*(10), 1338-U1163. doi: 10.1038/nn.2921

Aouizerate, B., Guehl, D., Cuny, E., Rougier, A., Bioulac, B., Tignol, J., & Burbaud, P. (2004). Pathophysiology of obsessive-compulsive disorder: A necessary link between phenomenology, neuropsychology, imagery and physiology. *Prog Neurobiol, 72*(3), 195-221. doi: 10.1016/j.pneurobio.2004.02.004

Bachevalier, J., & Loveland, K. A. (2006). The orbitofrontal-amygdala circuit and self-regulation of social-emotional behavior in autism. *Neuroscience and Biobehavioral Reviews, 30*(1), 97-117. doi: 10.1016/j.neubiorev.2005.07.002

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology, 37*(4-5), 407-419. doi: 10.1016/s0028-3908(98)00033-1

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., & Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron, 58*(4), 639-650. doi: 10.1016/j.neuron.2008.04.009

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 8*(12), 539-546. doi: 10.1016/j.tics.2004.10.003

Budhani, S., Marsh, A. A., Pine, D. S., & Blair, R. J. R. (2007). Neural correlates of response reversal: Considering acquisition. *Neuroimage, 34*(4), 1754-1765. doi: 10.1016/j.neuroimage.2006.08.060

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4*(6), 215-222. doi: 10.1016/s1364-6613(00)01483-2

Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., & Rosen, B. R. (2002). Dorsal anterior cingulate cortex: A role in reward-based decision making. *Proceedings of the National Academy of Sciences of the United States of America, 99*(1), 523-528. doi: 10.1073/pnas.012470999

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*(5364), 747-749. doi: 10.1126/science.280.5364.747

Chamberlain, S. R., Menzies, L., Hampshire, A., Suckling, J., Fineberg, N. A., del Campo, N., . . . Sahakian, B. J. (2008). Orbitofrontal dysfunction in patients with obsessive-compulsive disorder and their unaffected relatives. *Science, 321*(5887), 421-422. doi: 10.1126/science.1154433

Clatworthy, P. L., Lewis, S. J. G., Brichard, L., Hong, Y. T., Izquierdo, D., Clark, L., . . . Robbins, T. W. (2009). Dopamine release in dissociable striatal subregions predicts the different effects of oral methylphenidate on reversal learning and spatial working memory. *Journal of Neuroscience, 29*(15), 4690-4696. doi: 10.1523/jneurosci.3266-08.2009

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should i stay or should i go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B-Biological Sciences, 362*(1481), 933-942. doi: 10.1098/rstb.2007.2098

Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience, 22*(11), 4563-4567.

Cools, R., Lewis, S. J. G., Clark, L., Barker, R. A., & Robbins, T. W. (2007). L-dopa disrupts activity in the nucleus accumbens during reversal learning in parkinson's disease. *Neuropsychopharmacology, 32*(1), 180-189. doi: 10.1038/sj.npp.1301153

Cools, R., Sheridan, M., Jacobs, E., & D'Esposito, M. (2007). Impulsive personality predicts dopamine-dependent changes in frontostriatal activity during component processes of working memory. *Journal of Neuroscience, 27*(20), 5506-5514. doi: 10.1523/jneurosci.0601-07.2007

Dias, R., Robbins, T. W., & Roberts, A. C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the wisconsin card sort test: Restriction to novel situations and independence from ''on-line'' processing. *Journal of Neuroscience, 17*(23), 9285-9297.

Ersche, K. D., Roiser, J. P., Abbott, S., Craig, K. J., Muller, U., Suckling, J., . . . Bullmore, E. T. (2011). Response perseveration in stimulant dependence is associated with striatal dysfunction and can be ameliorated by a d-2/3 receptor agonist. *Biological Psychiatry, 70*(8), 754-762. doi: 10.1016/j.biopsych.2011.06.033

Evans, D. E., & Rothbart, M. K. (2007). Developing a model for adult temperament. *Journal of Research in Personality, 41*(4), 868-888. doi: 10.1016/j.jrp.2006.11.002

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *Neuroimage, 26*(2), 471-479. doi: 10.1016/j.neuroimage.2005.02.004

Freyer, T., Valerius, G., Kuelz, A. K., Speck, O., Glauche, V., Hull, M., & Voderholzer, U. (2009). Test-retest reliability of event-related functional mri in a probabilistic reversal learning task. *Psychiatry Research-Neuroimaging, 174*(1), 40-46. doi: 10.1016/j.pscychresns.2009.03.003

Grahn, J. A., Parkinson, J. A., & Owen, A. M. (2008). The cognitive functions of the caudate nucleus. *Progress in Neurobiology, 86*(3), 141-155. doi: 10.1016/j.pneurobio.2008.09.004

Guehl, D., Benazzouz, A., Aouizerate, B., Cuny, E., Rotge, J. Y., Rougier, A., . . . Burbaud, P. (2008). Neuronal correlates of obsessions in the caudate nucleus. *Biological Psychiatry, 63*(6), 557-562. doi: 10.1016/j.biopsych.2007.06.023

Hampshire, A., Chaudhry, A. M., Owen, A. M., & Roberts, A. C. (2012). Dissociable roles for lateral orbitofrontal cortex and lateral prefrontal cortex during preference driven reversal learning. *Neuroimage, 59*(4), 4102-4112. doi: 10.1016/j.neuroimage.2011.10.072

Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2006). The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *Journal of Neuroscience, 26*(32), 8360-8367. doi: 10.1523/jneurosci.1010-06.2006

Harrison, B. J., Soriano-Mas, C., Pujol, J., Ortiz, H., Lopez-Sola, M., Hernandez-Ribas, R., . . . Cardoner, N. (2009). Altered corticostriatal functional connectivity in obsessive-compulsive disorder. *Archives of General Psychiatry, 66*(11), 1189-1200.

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences, 16*(3), 174-180. doi: 10.1016/j.tics.2012.01.006

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science, 308*(5718), 78-83. doi: 10.1126/science.1108062

Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., & Homme, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: Characterization with rapid event-related fmri. *Neuroimage, 18*(2), 263-272. doi: 10.1016/s1053-8119(02)00057-5

Krugel, L. K., Biele, G., Mohr, P. N. C., Li, S. C., & Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Sciences of the United States of America, 106*(42), 17951-17956. doi: 10.1073/pnas.0905191106

Leh, S. E., Ptito, A., Chakravarty, M. M., & Strafella, A. P. (2007). Fronto-striatal connections in the human brain: A probabilistic diffusion tractography study. *Neuroscience Letters, 419*(2), 113-118. doi: 10.1016/j.neulet.2007.04.049

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fmri data sets. *Neuroimage, 19*(3), 1233-1239. doi: 10.1016/s1053-8119(03)00169-1

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306*(5695), 503-507.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu Rev Neurosci, 24*, 167-202. doi: 10.1146/annurev.neuro.24.1.167

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304*(5669), 452-454. doi: 10.1126/science.1094285

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature, 441*(7090), 223-226. doi: 10.1038/nature04676

Paulus, M. P., Feinstein, J. S., Leland, D., & Simmons, A. N. (2005). Superior temporal gyrus and insula provide response and outcome-dependent information during assessment and action selection in a decision-making situation. *Neuroimage, 25*(2), 607-615. doi: 10.1016/j.neuroimage.2004.12.055

Polli, F. E., Barton, J. J. S., Cain, M. S., Thakkar, K. N., Rauch, S. L., & Manoach, D. S. (2005). Rostral and dorsal anterior cingulate cortex make dissociable contributions during antisaccade error commission. *Proceedings of the National Academy of Sciences of the United States of America, 102*(43), 15700-15705. doi: 10.1073/pnas.0503657102

Robinson, O. J., Standing, H. R., DeVito, E. E., Cools, R., & Sahakian, B. J. (2010). Dopamine precursor depletion improves punishment prediction during reversal learning in healthy females but not males. *Psychopharmacology, 211*(2), 187-195. doi: 10.1007/s00213-010-1880-1

Rothbart, M. K., Ahadi, S. A., & Evans, D. E. (2000). Temperament and personality: Origins and outcomes. *Journal of Personality and Social Psychology, 78*(1), 122-135. doi: 10.1037//0022-3514.78.1.122

Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A pet investigation of agency. *Nature Neuroscience, 4*(5), 546-550.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*(5626), 1755-1758. doi: 10.1126/science.1082976

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593-1599. doi: 10.1126/science.275.5306.1593

Silvetti, M., Seurinck, R., & Verguts, T. (2011). Value and prediction error in medial frontal cortex: Integrating the single-unit and systems levels of analysis. *Frontiers in Human Neuroscience, 5*. doi: 10.3389/fnhum.2011.00075

Smillie, L. D., Cooper, A. J., Tharp, I. J., & Pelling, E. L. (2009). Individual differences in cognitive control: The role of psychoticism and working memory in set-shifting. *British Journal of Psychology, 100*, 629-643. doi: 10.1348/000712608x382094

Spreckelmeyer, Katja N., Krach, Soeren, Kohls, Gregor, Rademacher, Lena, Irmak, Arda, Konrad, Kerstin, . . . Gruender, Gerhard. (2009). Anticipation of monetary and social reward differently activates mesolimbic brain structures in men and women. *Social Cognitive and Affective Neuroscience, 4*(2), 158-165. doi: 10.1093/scan/nsn051

Tabibnia, G., Satpute, A. B., & Lieberman, M. D. (2008). The sunny side of fairness - preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science, 19*(4), 339-347. doi: 10.1111/j.1467-9280.2008.02091.x

Tekin, S., & Cummings, J. L. (2002). Frontal-subcortical neuronal circuits and clinical neuropsychiatry - an update. *Journal of Psychosomatic Research, 53*(2), 647-654. doi: 10.1016/s0022-3999(02)00428-2

Tricomi, E., Delgado, M. R., McCandliss, B. D., McClelland, J. L., & Fiez, J. A. (2006). Performance feedback drives caudate activation in a phonological learning task. *Journal of Cognitive Neuroscience, 18*(6), 1029-1043.

Tricomi, E., & Fiez, J. A. (2008). Feedback signals in the caudate reflect goal achievement on a declarative memory task. *Neuroimage, 41*(3), 1154-1167. doi: 10.1016/j.neuroimage.2008.02.066

Waegeman, A., Declerck, C. H., Boone, C., Van Hecke, W., & Parizel, P.M. . (2014). Individual differences in self-control in a time discounting task: An fmri study. *Journal of Neuroscience, Psychology, and Economics, 7*(2), 65-79. doi: 10.1037/npe0000018

Waltz, J. A., & Gold, J. M. (2007). Probabilistic reversal learning impairments in schizophrenia: Further evidence of orbitofrontal dysfunction. *Schizophrenia Research, 93*(1-3), 296-303. doi: 10.1016/j.schres.2007.03.010

Whiteside, S. P., Port, J. D., & Abramowitz, J. S. (2004). A meta-analysis of functional neuroimaging in obsessive-compulsive disorder. *Psychiatry Res, 132*(1), 69-79. doi: 10.1016/j.pscychresns.2004.07.001