

The Cerebellum Is Involved in Reward-based Reversal Learning

Patrizia Thoma · Christian Bellebaum · Benno Koch · Michael Schwarz · Irene Daum

Published online: 1 July 2008
© Springer Science + Business Media, LLC 2008

Abstract The cerebellum has recently been discussed in terms of a possible involvement in reward-based associative learning. To clarify the cerebellar contribution, eight patients with focal vascular lesions of the cerebellum and a group of 24 healthy subjects matched for age and IQ were compared on a range of different probabilistic outcome-based associative learning tasks, assessing acquisition, reversal, cognitive transfer, and generalization as well as the effect of reward magnitude. Cerebellar patients showed intact acquisition of stimulus contingencies, while reward-based reversal learning was significantly impaired. In addition, the patients showed slower acquisition of new stimulus contingencies in a second reward-based learning task, which might reflect reduced carry-over effects. Reward magnitude affected learning only during initial acquisition, with better learning on trials with high rewards in patients and control subjects. Overall, the findings suggest that the cerebellum is implicated in reversal learning as a dissociable component of reward-based learning.

Keywords Cerebellum · Reversal learning · Acquired equivalence · Associative learning · Reinforcement

Introduction

The detection, representation, and prediction of rewards, known to motivate goal-directed behavior, is primarily mediated by a dopaminergic circuitry involving the striatum and the prefrontal cortex (PFC) [1, 2]. Different components of reward-based learning seem to rely on distinct neural substrates. For instance, functional dissociations have been observed for the dorsal and ventral striatum: Reward-related activity has been shown in both parts, but modulations by reward magnitude and valence were only observed in the dorsal striatum [3, 4]. The latter also appears to be more involved in establishing links between actions and outcomes [5] and in learning from negative feedback, while the ventral striatum is more closely related to learning from positive feedback [3, 6, 7]. Additionally, a pattern of intact acquisition but impaired reversal of reward-based learning, which has been reported in medicated patients with Parkinson's disease and in patients with focal striatal lesions [8, 9], has been discussed in terms of dorsal/ventral striatal dissociations. Within the PFC, the ventral PFC has been related to reward-based reversal learning [10], the orbitofrontal cortex to the coding of reward magnitude [11] and of relative reward preference [12], and the anterior cingulate to the integration of reward, motor, and goal-relevant information [13].

In addition to fronto-striatal circuitry, the cerebellum might contribute to reward-based learning. It is interconnected with the PFC (areas 9/46 in particular) via reciprocal pathways, providing an anatomical basis for cerebellar mediation of non-motor "frontal" function [14–16]. Precise event timing might be one of the critical components coordinated by the cerebellum during learning [17]. The cerebellum has been implicated in different forms of motor and non-motor non-declarative learning [18, 19],

P. Thoma (✉) · C. Bellebaum · I. Daum
Institute of Cognitive Neuroscience,
Department of Neuropsychology, Faculty of Psychology,
Ruhr-University of Bochum,
44780 Bochum, Germany
e-mail: patrizia.thoma@rub.de

B. Koch · M. Schwarz
Department of Neurology,
Klinikum Dortmund,
44137 Dortmund, Germany

such as learning associations between stimuli [20, 21], sequence learning [22], classical eyeblink conditioning [23, 24], and probabilistic reasoning [25]. Cerebellar involvement has also been related to instrumental learning [26], and more recently, to reinforcement learning associated with substance abuse. Enhanced cerebellar activation was observed in substance abusers as they performed reward-learning tasks [27, 28], experienced drug craving [29, 30], and expected or recalled rewarding drug-related experiences [31, 32]. In healthy subjects, cerebellar activation was observed during the presentation of unpredicted rewards [33], the prediction of large future rewards as opposed to immediate rewards [34], and in association with rewarding sexual experiences [35].

According to current theories of reward processing, learning is based on prediction errors arising from an incongruity between expected and actually received rewards [36]. Prediction errors are coded within critical nodes of the reward–learning system, such as the striatum, the orbitofrontal cortex, and the anterior cingulate [e.g., 37], but also by the climbing fibers of cerebellar Purkinje cells. The function of prediction errors coded in the cerebellum was initially thought to be restricted to motor coordination on the basis of predicted sensory events [38], but was later extended to the cognitive domain, e.g., to the modification of internal thought and general action preparation [39]. It has also been suggested that the cerebellum completes roughly the same processing steps in different behavioral contexts (e.g., movement control, working memory, affective processing) [40].

Taken together, several lines of evidence suggest a cerebellar contribution to reward-based learning: (1) the involvement in various forms of associative learning, (2) frequent reward-related activations in neuroimaging studies, (3) the coding of sensorimotor prediction errors as well as (4) the notion of similar algorithms used by the cerebellum to process information in different contexts. Little is as yet known about the nature of the potential cerebellar involvement in reward-based associative learning. Thus, the present study investigated the effects of focal cerebellar damage on a broad range of different components of reward-based associative learning. Eight patients with vascular cerebellar lesions were compared to matched healthy controls. Two probabilistic associative learning tasks were administered, both assessing outcome-based learning of stimulus–stimulus associations. Although some of the previous studies rather investigated stimulus–outcome [5] or stimulus–response–outcome [13] associations, the present task did resemble procedures applied in some patient studies, where the outcome not only depended on the chosen stimulus, but also on the alternative not-chosen stimulus [41] or where the learning of an association between stimuli was based on feedback [42]. The first task

in the present study assessed both acquisition and reversal of stimulus–stimulus–outcome contingencies. As suggested by selective impairments observed in previous studies, distinct mechanisms might mediate reward-based acquisition and reversal. Performance during reversal might reflect adaptation of action strategies as well as processing of prediction errors [8, 43]. Different reward magnitudes were used to systematically vary the size of the prediction error, and because specific brain regions are concerned with the processing of reward magnitude (see above). If the cerebellum contributes to the prediction of reward and non-reward, prediction error processing might be disrupted in patients with cerebellar lesions.

Together with a categorization test following a second reversal at the end of the first task, the second learning task, an acquired equivalence paradigm, assessed the ability to transfer previously learned reward contingencies to new stimuli, showing clear parallels to the paradigm used by Myers et al. [42]. Generalization and cognitive transfer may be partly mediated by declarative memory strategies [42, 44], and the cerebellum might contribute to the formation of such strategies [45].

Methods

Subjects

Eight patients (three females, eight males) with selective vascular lesions of the cerebellum participated in this study. All were outpatients of the Klinikum Dortmund, assessed on average 60.4 months (SD 7.5) after the lesion event (range 48–73 months).

Lesions were documented by magnetic resonance imaging (MRI) using a standard three-dimensional T2-weighted sequence (1 mm×1 mm×5 mm voxel size). The affected brain regions were determined by visual inspection according to an established atlas [46] by two experienced independent raters. In three patients, only the right cerebellar hemisphere was affected (CERE3, CERE4, and CERE8), three patients suffered from unilateral left lesions (CERE5, CERE6, and CERE7), and two from bilateral lesions (CERE1 and CERE2). The deep nuclei were affected in two patients (CERE4 and CERE5) and nuclear damage could not be excluded in a third patient (CERE2). There was no MRI evidence of extracerebellar lesions in seven of the patients. One patient (CERE7) presented with a mild cerebral microangiopathy and minimal scar tissue on the left dorsolateral medulla oblongata. More detailed information on the cerebellar lesions is given in Table 1 and by the MR images in Fig. 1. Table 1 also includes information about the presence and severity of cerebellar signs, as determined during clinical neurological examina-

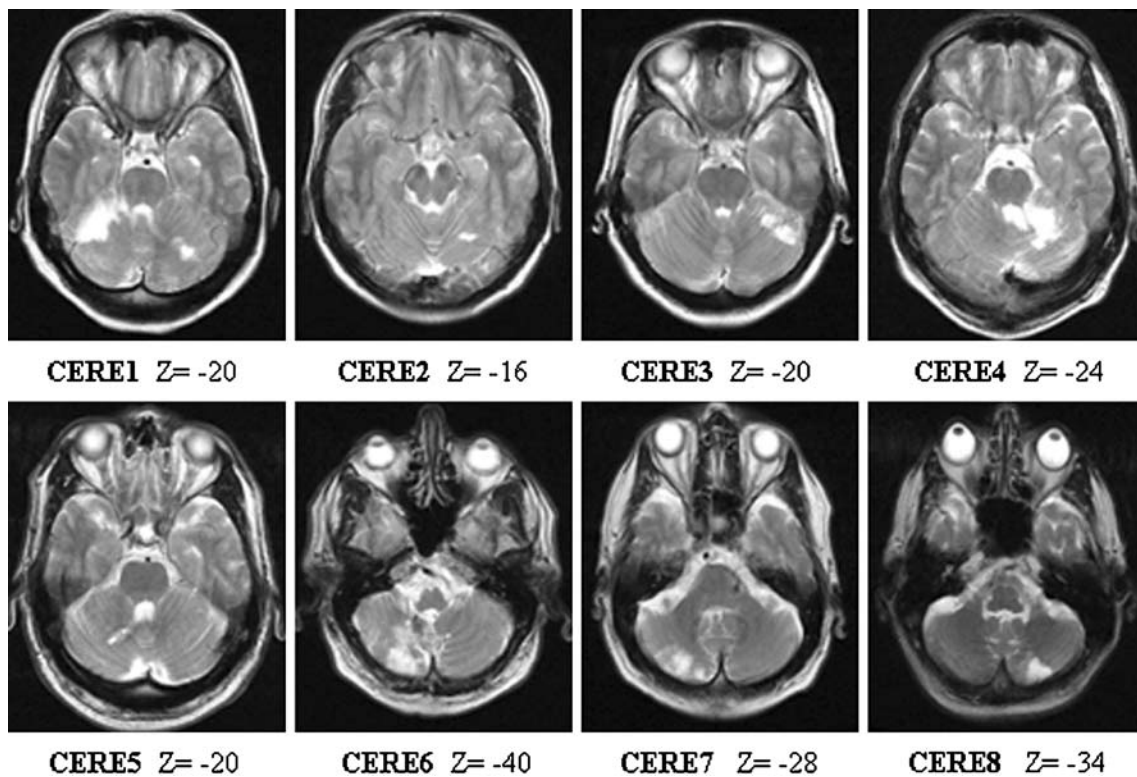


Fig. 1. T2-weighted transverse MR images of vascular lesion locations for the patients with cerebellar lesions (*CERE*)

tion. Overall motor and speech symptoms were very mild which indicates good recovery.

The control group comprised 24 healthy participants (ten males, 14 females), with three healthy controls being matched to each patient on age, IQ, and gender as closely as possible. Mean age was 54.8 years (SD 19.5) in the patient group (age range 30–79 years) and 53.7 years (SD 19.1) in the control group (age range 24–78 years). The mean IQ estimate, as assessed by the “Picture completion”

and “Similarities” subtests from a short German version of the Wechsler Adult Intelligence Scale [47], was 110.4 (SD 7.1) for the patients and 110.6 (SD 22.5) for the controls. Patients completed on average 11.1 years (SD 1.6) and controls 10.9 years (SD 2.2) of education. Mean age, mean estimated IQ, years of education, and the female/male ratio did not differ significantly between groups (all $p > 0.306$).

All subjects had normal or corrected-to-normal vision. Exclusion criteria for the controls were history of psychi-

Table 1 Overview of lesion sites and cerebellar signs in the cerebellar patients

Patient	LT	LS	Arterial territory				Cerebellar area				Cerebellar signs at the time of testing			
			AICA	ACS	PICA	AV	Anterior lobe	Posterior lobe	Vermis	Deep nuclei	Gait	Upper limb	Oculo-motor	Speech
CERE1	53	l, r	-	+	-	-	+,+	+,+	-,-	-,-	0	1	0	0
CERE2	64	l, r	-	+	-	-	-,+	+,+	-,-	Cannot be excluded	0	1	0	1
CERE3	62	r	+	+	-	-	+	+	-	-	0	1	0	0
CERE4	62	r	-	+	-	+	+	+	-	+	1	1	0	1
CERE5	73	l	-	+	-	-	+	+	-	+, probably dentate	0	1	0	1
CERE6	48	l	-	-	+	-	-	+	-	-	0	0	0	0
CERE7	62	l	-	-	+	-	-	+	-	-	0	0	0	0
CERE8	59	r	+	-	-	-	-	+	-	-	1	0	0	0

Scoring of cerebellar signs on a four-point rating scale: 0 = normal, 1 = slight, 2 = moderate, 3 = severe

LT Time since lesion event (months), LS lesion side (l left, r right), AICA arteria cerebelli inferior anterior, ACS arteria cerebelli superior, PICA arteria cerebelli inferior posterior, AV arteria vertebralis

atric or neurological disorders and regular use of medication affecting the central nervous system. Apart from their cerebellar lesions, the patients did not present with any neurological disorders.

The study conforms with the Declaration of Helsinki and was approved by the local ethics board, and all subjects gave written informed consent prior to the assessment. Subjects were reimbursed for their participation with a minimum of 20 euro. This sum could be increased depending on the subjects' performance in the reward learning tasks.

The test battery (~2 h) entailed a brief screening of IQ and declarative learning abilities, followed by two reward-based learning tasks and a brief screening of executive function.

Cognitive Screening

Present-state IQ was estimated by means of the “Picture completion” and “Similarities” subtests from the German short version of the Wechsler Adult Intelligence Scale [47]. The immediate and delayed recall subtests of the verbal paired associates task from the Wechsler Memory Scale—Revised [48] were used to assess the ability to learn declaratively across trials. In this task, eight word pairs were read out to the subject. Thereafter, the first word of each pair was read out and subjects had to reproduce the second word from memory. Half of the pairs (“easy pairs”) contained semantically related words (e.g., “metal/iron”

and the other half (“hard pairs”) unrelated words (e.g., “salad/pen”). Learning was completed when all pairs were correctly reproduced (min. three trials; max. six trials), and analysis of immediate recall was based on the first three trials only. Delayed recall was assessed after 20 min. Due to time restrictions, verbal paired associate learning could not be assessed in two control subjects.

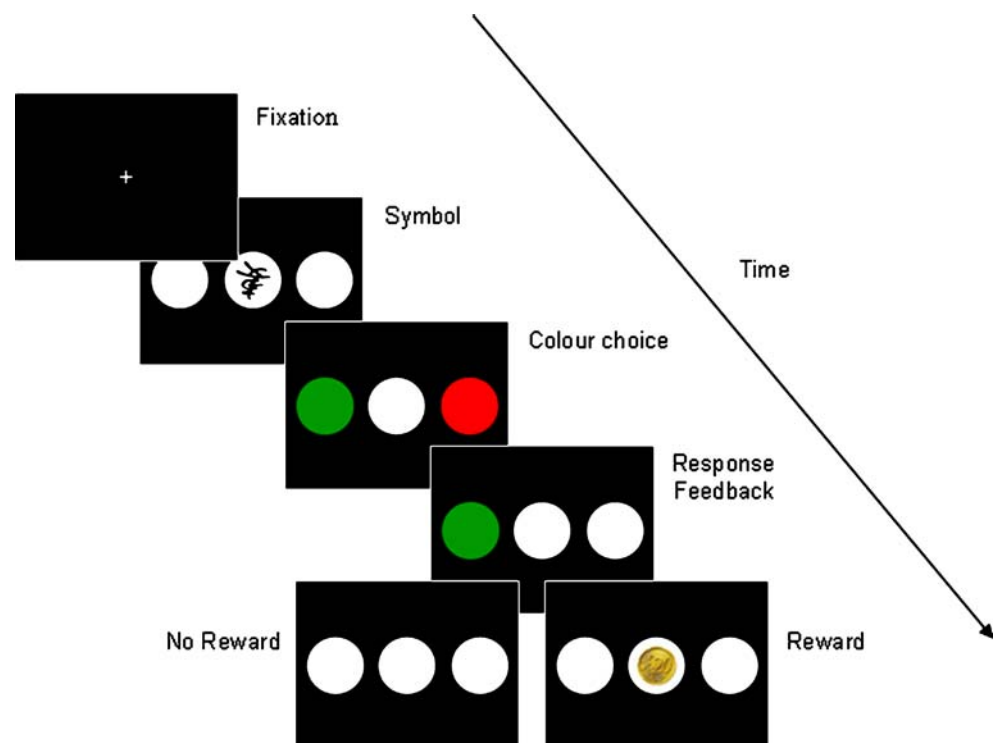
Reward-based Learning Tasks

Two probabilistic learning tasks (see [8, 43] for a detailed description) were used to assess different components of reward-based learning.

Acquisition and Reversal of Reward-based Associative Learning

Subjects had to learn associations between two colors (red or green) and four Asian symbols on the basis of feedback given after each response. Each trial (see Fig. 2) started with a fixation cross in the center of the screen, followed by the randomized presentation of one of the Asian symbols and then a black screen showing a red, a white, and a green circle horizontally aligned (white circle in the center, randomized locations of the red and the green circles). Subjects were instructed to select either the red or the green circle by pressing one of two response buttons and to try to find out how to respond correctly on the basis of the feedback they would receive. They were informed before

Fig. 2. Task design for the reward learning paradigms



participation that they would be paid the sum of all their rewards in the end. The color of the chosen circle changed to white after the button press and feedback about the response was provided: Choosing the correct color associated with the symbol in question was rewarded (display of a 5 cent or 20 cent coin at the center of the screen) with a probability of 80%, i.e., not every correct response was reinforced in order to prevent rapid insight into stimulus–stimulus–outcome associations and to reduce the contribution of declarative learning strategies. Rewarded and non-rewarded trials were counterbalanced across the experiment and symbol-to-color pairings were counterbalanced across subjects. Throughout the task, two symbols were associated with a reward of 5 cents and the remaining two with 20 cents. On non-reinforced trials or on trials with incorrect responses, three empty white circles appeared.

The task comprised three stages: During acquisition (three blocks, 40 trials each), two of the four Asian symbols were associated with the color green and the remaining two with red. A reversal phase (three blocks, 40 trials each) followed, in which these symbol-to-color pairings were reversed without notice. The third stage (“single symbol reversal”) involved learning trials and test trials. During learning, the symbol-to-color-associations were switched back to the pairings of initial acquisition, again without notice. However, only two of the four familiar symbols were presented (one associated with green, one with red). When the subjects had learned the new reversal—as indicated by five correct successive responses in a row (min. 15 trials, max. 50 trials)—a test phase started (40 trials), involving all four symbols again. To respond correctly, subjects had to transfer the reversed associations to the two symbols that were not used during learning. Trial-to-trial feedback was no longer provided, but subjects were informed about the sum of their winnings every five trials to keep them motivated.

Acquired Equivalence Learning

The second probabilistic learning task entailed the same trial structure as the first task (see Fig. 2), the same reward probability (80%), but associations between four new Asian symbols and two new colors (pink and brown) had to be learned and the reward magnitude was kept constant (5 cents).

Table 3 Symbol–color pairings for the three phases of the acquired equivalence task

Acquisition phase 1	Acquisition phase 2	Test phase
Symbol 1—pink	Symbol 1—blue	Symbol 1—blue
Symbol 2—pink		Symbol 2—blue
Symbol 3—brown	Symbol 3—yellow	Symbol 3—yellow
Symbol 4—brown		Symbol 4—yellow

This task again involved three stages: Initial acquisition of the new symbol-to-color pairings was completed when subjects reached eight correct responses in a row (min. 38 trials, max. 80 trials). During the subsequent second acquisition, only two of the four symbols (one initially associated with pink, one with brown) were presented, and each one was now associated with one of two new colors (blue and yellow). This learning stage was terminated after five correct responses in a row were reached (min. 15 trials, max. 80 trials) and followed by a test phase, comparable to the one in the single symbol reversal task (40 trials, no trial-to-trial feedback, information about cumulative winnings every five trials). During test trials, subjects had to transfer their knowledge about the new symbol-to-color pairings to the other two symbols (used for initial learning) by applying the rule that the two symbols were equivalent in the sense that they were always associated with the same color. Tables 2 and 3 provide a detailed overview of the symbol-to-color pairings in the different stages of the reward-based learning tasks.

Assessment of Executive Function

A computerized version of the modified card sorting test (MCST) [49] was administered to assess cognitive flexibility in the patients during clinical screening (data are reported relative to the performance of 12 age-matched healthy controls). Subjects were asked to sort 48 test cards according to four stimulus cards each of which was unique in terms of its color, shape, and number of items. Participants were instructed to find out the current sorting rule based on the feedback given after each trial. A change of the sorting principle was announced after every six correct responses and the new rule had to be found. The

Table 2 Symbol–color pairings for the three phases of reward learning task 1

Acquisition	Reversal	Single symbol reversal	
		Acquisition	Test
Symbol 1—red	Symbol 1—green	Symbol 1—red	Symbol 1—red
Symbol 2—red	Symbol 2—green		Symbol 2—red
Symbol 3—green	Symbol 3—red	Symbol 3—green	Symbol 3—green
Symbol 4—green	Symbol 4—red		Symbol 4—green

number of completed categories, perseverative, and non-perseverative errors was recorded.

Statistical Analyses

Statistical analyses were performed using SPSS 15.0 software. Patient and control groups were compared using (repeated-measures) analyses of variance (ANOVAs) or *t* tests where appropriate. When the conditions for parametric testing were not fulfilled, Mann–Whitney *U* tests were used. The significance level was set at 0.05, two-tailed.

Results

Verbal Paired Associates

Verbal paired associates data are presented in Fig. 3a (easy pairs) and b (hard pairs). Repeated-measures ANOVA of the immediate recall performance with factors group (patients vs. controls), item difficulty (easy vs. hard), and block (one to three) yielded significant main effects of group ($F_{1, 28}=5.717$; $p=0.024$), difficulty ($F_{1, 28}=33.923$; $p<0.001$), and

block (linear trend $F_{1, 28}=52.905$; $p<0.001$), a significant block \times difficulty interaction (linear trend $F_{1, 28}=26.093$; $p<0.001$), a marginally significant difficulty \times group interaction ($F_{1, 28}=4.167$; $p=0.051$), and a marginally significant three-way interaction (linear trend $F_{1, 28}=3.267$; $p=0.081$). No other effects reached significance (all $p>0.173$). The main effects indicate a better overall performance in controls relative to patients, generally better performance for easy vs. hard pairs and significant learning across blocks. The block \times difficulty interaction reflects a larger performance difference between easy and hard items in the first block compared to later blocks. The marginally significant difficulty \times group interaction reveals overall better performance of controls relative to patients on difficult items ($F_{1, 28}=5.717$; $p=0.024$), but not on easy items ($p=0.303$), while the marginally significant three-way interaction reflects a larger difference between easy and hard items in the patient group at the beginning of learning which decreases across blocks (see Fig. 3). There was no significant group difference for delayed recall of easy or hard pairs after 20 min (both $p>0.193$).

Reward-based Learning Tasks

Acquisition and Reversal

Learning curves for acquisition and reversal tasks are presented in Fig. 4. Repeated-measures ANOVAs were performed separately for acquisition and reversal, with group (patients vs. controls), block (three blocks of 40 trials each), and reward (5 vs. 20 cents) as factors. Analysis of the percentage of correct responses during acquisition yielded a significant main block effect (linear trend $F_{1, 30}=23.195$; $p<0.001$) and a marginally significant block \times reward interaction (linear trend $F_{1, 30}=3.330$; $p=0.078$). The block effect reflects significant learning, with a trend towards better learning with 20 cent rewards (linear trend $F_{1, 30}=25.153$; $p<0.001$) compared to 5 cent rewards (linear trend $F_{1, 30}=3.953$; $p=0.056$). No other effects approached significance (all $p>0.340$).

Analysis of reversal yielded a trend towards a block effect (linear trend $F_{1, 30}=3.379$; $p=0.076$), reflecting marginal learning across blocks, and a significant group effect ($F_{1, 30}=4.582$; $p=0.041$), with a lower percentage of correct responses of the patients relative to the control group. No other effects approached significance (all $p>0.133$).

Single Symbol Reversal

Repeated-measures ANOVA of the percentage of correct responses (see Fig. 5a) in the transfer phase involving the factors symbol type (learned vs. transfer) and group

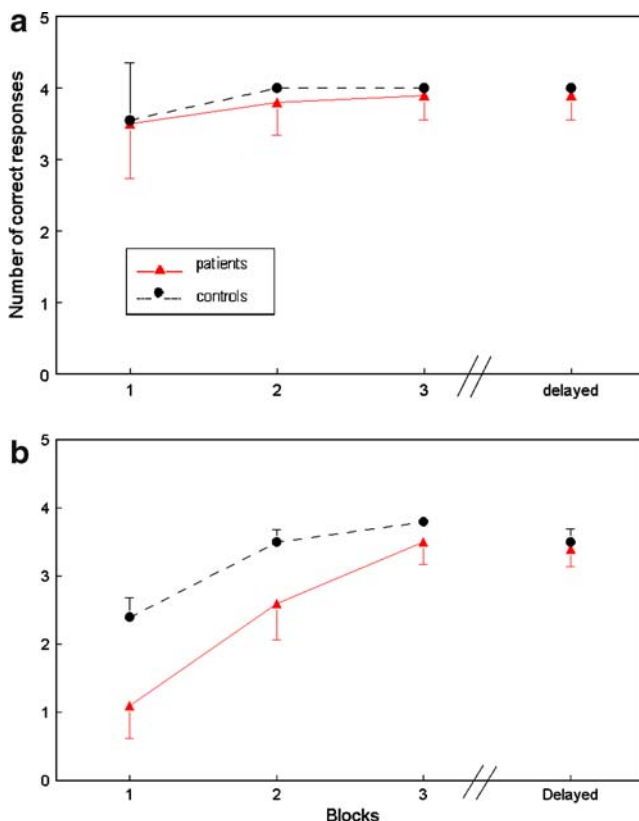
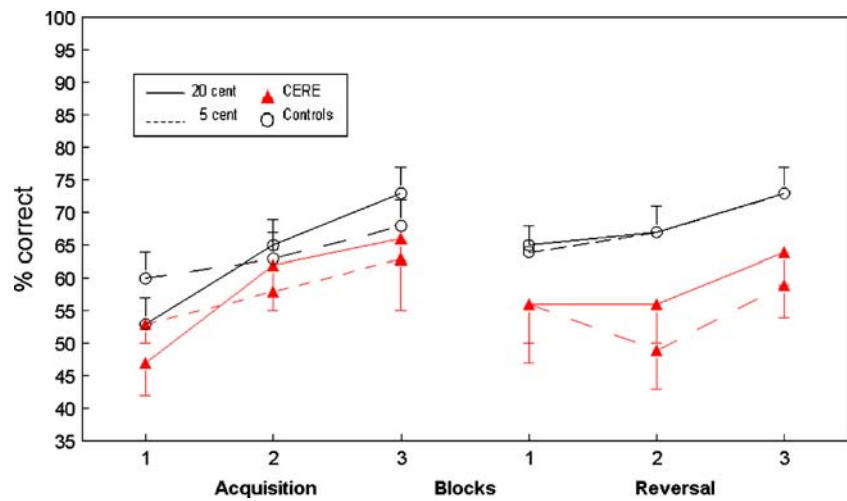


Fig. 3. Learning curves for patients and controls in the paired associates task. Error bars represent standard error of the mean. **a** Easy items. **b** Hard items

Fig. 4. Learning curves for the acquisition and the reversal phase of reward learning Task 1 for CERE patients and controls. Error bars represent standard error of the mean



(patients vs. controls) yielded a marginally significant main effect of symbol type ($F_{1, 30}=3.702$; $p=0.064$), with better performance with learned compared to transfer symbols. The ANOVA did not yield a significant effect of group and no significant group \times symbol-type interaction (both $p>0.323$).

Acquired Equivalence

The acquired equivalence task required transfer of stimulus contingencies across the study material and could not be accomplished by reference to earlier stimulus contingencies.

In a first step, we analyzed how fast the subjects learned the new stimulus–response contingencies in the first acquisition phase of the acquired equivalence task. Patients and controls did not differ in terms of the trials to criterion taken to achieve the learning criterion of eight successive correct responses when *all* subjects were considered (medians and absolute deviations: patients 62.0 trials \pm 11.1/controls 48.0 trials \pm 16.1/ $p=0.319$). Interestingly, the patients took significantly longer to reach the learning criterion, when *only* learners were considered in both groups, i.e., those subjects who reached the learning criterion at all within 80 trials (16 control subjects and seven patients; medians and absolute deviations: patients 60.0 trials \pm 9.7/controls 41.5 trials \pm 6.1/ $U=13.500$; $p=0.004$).

Performance in the test phase (see Fig. 5b) was analyzed by repeated-measures ANOVA of the percentage of correct responses, with factors group (all patients vs. all controls, i.e., considering learners *and* non-learners in both groups) and symbol type (learned vs. transfer) and yielded a significant symbol-type effect ($F_{1, 30}=6.238$; $p=0.018$), indicating better performance with learned than with transfer symbols. There were no significant main group or interaction effects (both $p>0.332$).

Executive Function

As can be inferred from Table 4, CERE patients did not differ from a matched healthy reference group with regard

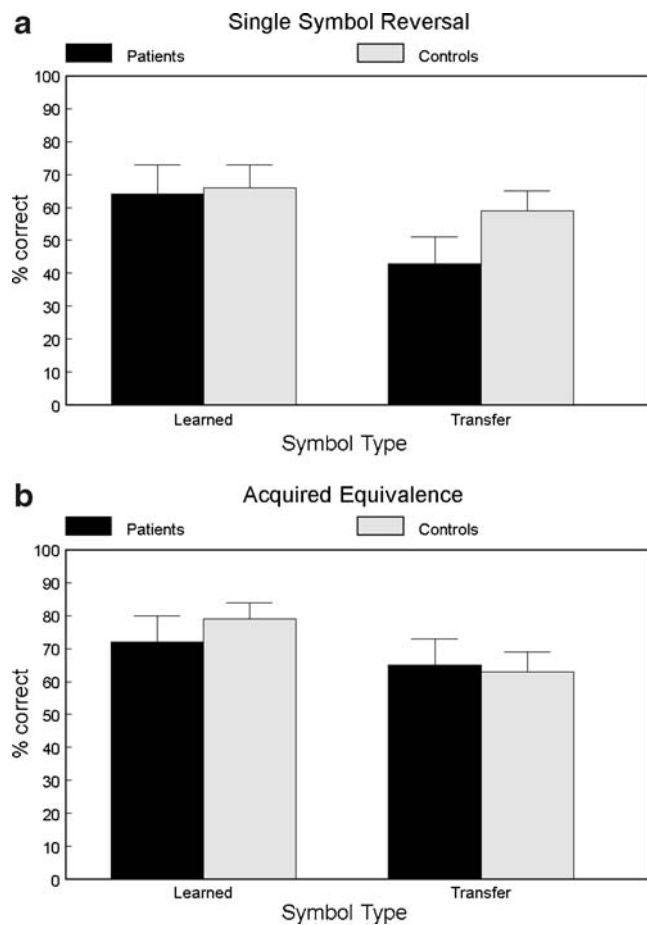


Fig. 5. Results from the test phases of the single symbol reversal and the acquired equivalence task test phases. Error bars represent standard errors of the mean. **a** Single symbol reversal task. **b** Acquired equivalence

Table 4 Performance of the cerebellar patients (CERE) on the modified card sorting test (MCST) as compared to a matched healthy reference group that did not participate in the present study

	CERE patients	Controls	p-value
<i>N</i>	8	12	
Mean age	50.00 (19.89)	56.33 (12.45)	0.390
Mean estimated IQ	110.56 (7.22)	110.42 (6.93)	0.964
MCST			
No. of categories completed	4.63 (1.41)	4.75 (1.66)	0.863
No. of perseverations	3.63 (4.00)	3.08 (4.12)	0.774
No. of non-perseverative errors	8.63 (4.41)	7.92 (5.28)	0.758

to the number of completed categories, perseverations, or non-perseverative errors in the MCST (all $p > 0.758$).

Discussion

The findings of this study indicate a selective impairment of reward-based reversal learning after focal cerebellar damage. Cerebellar patients initially acquired stimulus–stimulus–outcome associations on a level comparable to matched healthy controls but showed a clear impairment during reversal. During acquisition, there was a general trend for better learning with higher compared to smaller rewards. Although reward magnitude did not significantly affect performance in the reversal phase, the acquisition effect supports the idea of a true involvement of reward-based learning processes as opposed to pure feedback-learning accounts of the present task. There was no significant difference between groups with respect to transfer performance following an additional reversal. In the acquired equivalence task, which required stimulus generalization and transfer, acquisition of new stimulus–stimulus–outcome contingencies was markedly slowed in the patients, when only subjects defined as “learners” were considered in the analyses. Again, patients and controls did not differ in terms of response accuracy during the stimulus generalization and transfer phase of the acquired equivalence task.

The following sections address the patients’ impairments in the first reversal phase, the acquisition deficits in the acquired equivalence task, and alternative interpretations of the observed deficit pattern.

Impairment of Reward-based Reversal Learning

The most remarkable finding in the present study was impaired reward-based reversal learning in the cerebellar lesion group, a deficit which has previously been documented in rodents with cerebellar damage [50]. The present data extend previous findings which related reward-based reversal learning to the ventral PFC [10] and the ventral [9] and dorsal striatum [8]. In particular, it has been suggested that reward-based reversal learning depends on dopaminergic striatal fine-tuning [51]. The cerebellum might have an

additional modulatory role in this regard, an idea which is also supported by evidence that cerebellar contributions to reward processing seem to manifest themselves most clearly when the function of other nodes in the reward circuitry is compromised. For instance, enhanced cerebellar activity during reward processing has been observed in patients suffering from Parkinson’s disease [52] or attention deficit hyperactivity disorder [53], but not in healthy control subjects. Interestingly, Bellebaum et al. [8] recently reported a similar pattern of deficits, characterized by impaired reward-based reversal performance and reduced carry-over effects in later acquisition stages, in patients with focal basal ganglia lesions. It is conceivable that the cerebellum is primarily concerned with the coordination and timing of reward-based reversal learning processes in other parts of the reward circuitry.

The cerebellum has been shown to be involved in coding sensorimotor prediction errors [54]. Lesion-induced attenuation of prediction error signaling might thus have reduced the patients’ ability to efficiently adapt to changing stimulus contingencies in the reversal learning conditions. Ramnani et al. [33] suggested that the cerebellar vermis might be involved in processing affective information related to the unexpected occurrence of rewards. It is unlikely that prediction error signaling was completely absent in the patients, as marginally significantly better learning with higher rewards occurred in the subjects, and better learning has been related to increased prediction errors [e.g., 55]. The reward magnitude modulation of performance in the patients might have been mediated by intact striatal, dopaminergic systems, thought to code longer-lasting prediction errors with a globally alerting effect supporting the long-term biasing, prioritizing, and selection of some action over others. Cerebellar Purkinje cells, on the other hand, are thought to form part of an anatomically more selective system, concerned with providing immediate positive or negative corrective feedback allowing for particularly fast behavioral changes [39, 56]. O’Doherty et al. [5] proposed, in accordance with the actor-critic model of reward-based learning [57], that learning the values of an action and their implementation via a certain response policy might rely on dissociable neural substrates within the striatum. However, the change of reaction policy necessary

at the beginning of reward-based reversal might particularly draw upon cerebellar coding of error prediction/negative feedback [58]. This interpretation is in accordance with evidence that reward learning based on positive and negative feedback is mediated by distinct neural substrates, up to now mainly discussed in terms of dorsal/ventral striatal dissociations [7, 59]. It is plausible that reversal learning of acquired stimulus–stimulus–outcome associations might at least initially rely more on negative feedback (non-reward of previously rewarded associations) than on positive feedback (reward of new associations). However, further studies are needed to explicitly address the question whether the cerebellum might be differentially implicated in learning from positive and negative feedback in reward learning.

Acquisition of Rewarded Stimulus–Stimulus Associations—Impaired or Not?

It is interesting that cerebellar patients acquired the initial contingencies in the first task at a similar level as controls, but were clearly slower during the acquisition of new stimulus contingencies in the acquired equivalence paradigm. This was the case, although the proportion of non-learners in the latter task was even higher in the controls than in the patient group (33.3% relative to 12.5%). Controls may have used their experience with the first acquisition task in learning another reward-based task, thereby showing transfer or generalization effects, while the patients may not have benefited from experience in a comparable way, possibly due to attenuated cerebellar prediction error signaling.

The performance differences cannot be explained by the different methods of analyzing the data from the two acquisition phases (“percentage correct” only for the first acquisition and “trials to criterion” for the acquired equivalence acquisition). When an additional analysis of “trials to criterion” (eight consecutive correct responses) was performed for the acquisition phase of the first task in analogy to the acquired equivalence task phase 1 analyses, there was still no indication of a group difference ($p=0.750$) on acquisition speed during the initial acquisition phase in the first task.

Alternative Interpretations of the Observed Impairment Pattern?

The patients were assessed on average 5 years after the occurrence of the lesion event, so that acute effects such as edema or severe motor impairments did not interfere with learning. There were no significant group differences on reaction times ($p=0.154$), which also did not correlate with the degree of reward-based learning, as reflected by correct

responses in the acquisition, reversal, and acquired equivalence subtasks and by trials-to-criterion in the latter task (all $p>0.217$). As can be inferred from Table 1, cerebellar signs in the patients were very mild—if present at all—at the time of participation, and none of the patients rated these residual symptoms as significant for activities of daily life. It is possible that some degree of functional reorganization may have occurred since the lesion event.

An alternative explanation of the reward-based reversal deficit in the cerebellar lesion patients would imply a more general associative learning deficit—irrespective of the involvement of reward. In a previous study [21], cerebellar patients have been found to be impaired in learning stimulus–stimulus associations (colors and numbers). However, these patients suffered from degenerative cerebellar disorders, while the patients in the present study presented with very focal lesions of the cerebellum, mainly restricted to the cerebellar cortex. Also, the comparable performance of patients and controls during acquisition of the reward tasks argues against a general deficit in learning stimulus–stimulus associations because it is unlikely that such a problem would manifest itself only during reversal learning.

It is also unlikely that the observed reversal learning deficit can be fully attributed to a disruption of cerebellar output to the PFC or to frontally mediated executive deficits: Cerebellar efferent connections mainly target the dorsolateral PFC [60], but the ventral/orbitofrontal PFC has been linked to reversal learning [61], and there was no evidence of an executive impairment in the current sample of cerebellar patients (see MCST data).

Alternatively, disrupted interaction between cerebellar and midbrain structures might be considered to underlie the observed deficits in the patients, mediated by the reciprocal connections, part of them dopaminergic, between the cerebellar vermis and the ventral tegmental area [62]. However, none of the patients presented with a vermis lesion in the present study, supporting the notion that damage to the cerebellum itself and not disrupted communication with the midbrain underlies the impairment pattern.

Lesions were restricted to the cerebellum in all but one patient. When the data of the only patient with additional mild cerebral microangiopathy and minimal brainstem damage were analyzed on a single case basis, he was not found to be significantly impaired on any of the reward-based learning paradigms. Thus, it is unlikely that the extracerebellar lesions in this patient affected the overall impairment pattern of the cerebellar group.

Given the probabilistic nature of the present tasks and the fact that the stimuli were difficult to verbalize and semantically unrelated, non-declarative learning was assumed to play a major role. However, generalization and transfer learning additionally appear to rely on declarative strategies mediated by the medial temporal lobe [42, 44],

and there was evidence of impaired initial declarative learning of semantically unrelated word pairs, as assessed by the paired associates task, in the cerebellar patients. According to post-experimental interviews, none of the patients gained full insight into stimulus contingencies, suggesting an inability to efficiently use declarative strategies. It is noteworthy, however, that most control subjects did not gain full insight into task contingencies, either and that, with only 59% and 63% correct responses in the transfer conditions, their performance was rather poor, suggesting that declarative mechanisms might not have been used efficiently in either subject group, probably due to the probabilistic nature of the task. In addition to this, it is unlikely that impaired declarative learning might have contributed to the reward-based reversal learning impairment in the patients, since there is no evidence from the literature that declarative learning might play a significant role during probabilistic reversal learning.

Conclusions and Future Directions

The findings of the present study indicate that in humans, even small cerebellar lesions, which were restricted to the cerebellar cortex in a subgroup of subjects, affect reversal components of reward-based associative learning. Furthermore, the present data support the notion of reversal learning as a dissociable component of reward-based learning which may be selectively affected despite intact acquisition.

Studies involving larger samples of cerebellar patients with differing lesion locations are clearly needed to extend our sparse knowledge about the contributions of different cerebellar regions to different aspects of reward-based learning.

Acknowledgment We gratefully acknowledge support by the German Ministry for Education and Research (Bundesministerium für Bildung und Forschung, BMBF, Grant Nr. 01 GW0541).

References

1. Delgado MR (2007) Reward-related responses in the human striatum. *Ann N Y Acad Sci* 1104:70–88
2. Schultz W (2007) Behavioral dopamine signals. *Trends Neurosci* 30(5):203–210
3. Delgado MR, Locke HM, Stenger VA, Fiez JA (2003) Dorsal striatum responses to reward and punishment: effects of valence and magnitude manipulations. *Cogn Affect Behav Neurosci* 3(1):27–38
4. Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA (2000) Tracking the hemodynamic responses to reward and punishment in the striatum. *J Neurophysiol* 84(6):3072–3077
5. O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304(5669):452–454
6. Yacubian J, Glascher J, Schroeder K, Sommer T, Braus DF, Buchel C (2006) Dissociable systems for gain- and loss-related

- value predictions and errors of prediction in the human brain. *J Neurosci* 26(37):9530–9537
7. Seymour B, Daw N, Dayan P, Singer T, Dolan R (2007) Differential encoding of losses and gains in the human striatum. *J Neurosci* 27(18):4826–4831
8. Bellebaum C, Koch B, Schwarz M, Daum I (2008) Focal basal ganglia lesions are associated with impairments in reward based reversal learning. *Brain* 131:829–841
9. Cools R, Barker RA, Sahakian BJ, Robbins TW (2001) Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cereb Cortex* 11(12):1136–1143
10. Cools R, Clark L, Owen AM, Robbins TW (2002) Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J Neurosci* 22(11):4563–4567
11. O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* 4(1):95–102
12. Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. *Nature* 398(6729):704–708
13. Holroyd CB, Coles MG (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 109(4):679–709
14. Daum I, Ackermann H (1997) Neuropsychological abnormalities in cerebellar syndromes—fact or fiction? *Int Rev Neurobiol* 41:455–471
15. Schmahmann JD (2004) Disorders of the cerebellum: ataxia, dysmetria of thought, and the cerebellar cognitive affective syndrome. *J Neuropsychiatry Clin Neurosci* 16(3):367–378
16. Timmann D, Daum I (2007) Cerebellar contributions to cognitive functions: a progress report after two decades of research. *Cerebellum* 6(3):159–162
17. Ivry RB, Spencer RM, Zelaznik HN, Diedrichsen J (2002) The cerebellum and event timing. *Ann N Y Acad Sci* 978:302–317
18. Leggio MG, Neri P, Graziano A, Mandolesi L, Molinari M, Petrosini L (1999) Cerebellar contribution to spatial event processing: characterization of procedural learning. *Exp Brain Res* 127(1):1–11
19. Raymond JL, Lisberger SG, Mauk MD (1996) The cerebellum: a neuronal learning machine? *Science* 272(5265):1126–1131
20. Timmann D, Drepper J, Calabrese S, Burgerhoff K, Maschke M, Kolb FP et al (2004) Use of sequence information in associative learning in control subjects and cerebellar patients. *Cerebellum* 3(2):75–82
21. Drepper J, Timmann D, Kolb FP, Diener HC (1999) Non-motor associative learning in patients with isolated degenerative cerebellar disease. *Brain* 122(Pt 1):87–97
22. Doyon J, Song AW, Karni A, Lalonde F, Adams MM, Ungerleider LG (2002) Experience-dependent changes in cerebellar contributions to motor sequence learning. *Proc Natl Acad Sci U S A* 99(2):1017–1022
23. Daum I, Schugens MM, Ackermann H, Lutzenberger W, Dichgans J, Birbaumer N (1993) Classical conditioning after cerebellar lesions in humans. *Behav Neurosci* 107(5):748–756
24. Gerwig M, Kolb FP, Timmann D (2007) The involvement of the human cerebellum in eyeblink conditioning. *Cerebellum* 6(1):38–57
25. Blackwood N, Ffytche D, Simmons A, Bentall R, Murray R, Howard R (2004) The cerebellum and decision making under uncertainty. *Brain Res Cogn Brain Res* 20(1):46–53
26. Lalonde R (1994) Cerebellar contributions to instrumental learning. *Neurosci Biobehav Rev* 18(2):161–170
27. Anderson CM, Maas LC, Frederick B, Bendor JT, Spencer TJ, Livni E et al (2006) Cerebellar vermis involvement in cocaine-related behaviors. *Neuropsychopharmacology* 31(6):1318–1326

28. Martin-Solch C, Magyar S, Kunig G, Missimer J, Schultz W, Leenders KL (2001) Changes in brain activation associated with reward processing in smokers and nonsmokers. A positron emission tomography study. *Exp Brain Res* 139(3):278–286
29. Olbrich HM, Valerius G, Paris C, Hagenbuch F, Ebert D, Juengling FD (2006) Brain activation during craving for alcohol measured by positron emission tomography. *Aust N Z J Psychiatry* 40(2):171–178
30. Bonson KR, Grant SJ, Contoreggi CS, Links JM, Metcalfe J, Weyl HL et al (2002) Neural systems and cue-induced cocaine craving. *Neuropsychopharmacology* 26(3):376–386
31. Volkow ND, Wang GJ, Ma Y, Fowler JS, Zhu W, Maynard L et al (2003) Expectation enhances the regional brain metabolic and the reinforcing effects of stimulants in cocaine abusers. *J Neurosci* 23(36):11461–11468
32. Grant S, London ED, Newlin DB, Villemagne VL, Liu X, Contoreggi C et al (1996) Activation of memory circuits during cue-elicited cocaine craving. *Proc Natl Acad Sci U S A* 93(21):12040–12045
33. Ramnani N, Elliott R, Athwal BS, Passingham RE (2004) Prediction error for free monetary reward in the human prefrontal cortex. *Neuroimage* 23(3):777–786
34. Tanaka SC, Doya K, Okada G, Ueda K, Okamoto Y, Yamawaki S (2004) Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat Neurosci* 7(8):887–893
35. Holstege G, Georgiadis JR (2004) The emotional brain: neural correlates of cat sexual behavior and human male ejaculation. *Prog Brain Res* 143:39–45
36. Schultz W (1998) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80(1):1–27
37. Bellebaum C, Daum I (2008) Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *Eur J Neurosci* 27(7):1823–1835
38. Marr D (1969) A theory of cerebellar cortex. *J Physiol* 202(2):437–470
39. Schultz W, Dickinson A (2000) Neuronal coding of prediction errors. *Annu Rev Neurosci* 23:473–500
40. Ramnani N (2006) The primate cortico-cerebellar system: anatomy and function. *Nat Rev Neurosci* 7(7):511–522
41. Frank MJ, Seeberger LC, O’D’Eilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306(5703):1940–1943
42. Myers CE, Shohamy D, Gluck MA, Grossman S, Kluger A, Ferris S et al (2003) Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *J Cogn Neurosci* 15(2):185–193
43. Weiler JA, Bellebaum C, Daum I (2008) Aging affects acquisition and reversal of reward-based associative learning. *Learn Mem* 15(4):190–197
44. Keri S, Nagy O, Kelemen O, Myers CE, Gluck MA (2005) Dissociation between medial temporal lobe and basal ganglia memory systems in schizophrenia. *Schizophr Res* 77(2–3):321–328
45. Weis S, Klaver P, Reul J, Elger CE, Fernandez G (2004) Temporal and cerebellar brain regions that support both declarative memory formation and retrieval. *Cereb Cortex* 14(3):256–267
46. Matsui T, Hiranu A (1978) An atlas of the human brain for computerized tomography. Fischer, Stuttgart
47. Dahl G (1972) Reduzierter Wechsler Intelligenztest (short version of the Wechsler Intelligence Scale). Hain, Meisenheim
48. Wechsler D (1981) The Wechsler Adult Intelligence Scale—revised. The Psychological Corporation, San Antonio
49. Nelson HE (1976) A modified card sorting test sensitive to frontal lobe defects. *Cortex* 12:313–324
50. Stanton ME, Peloso E, Brown KL, Rodier P (2007) Discrimination learning and reversal of the conditioned eyeblink reflex in a rodent model of autism. *Behav Brain Res* 176(1):133–140
51. Robinson S, Sandstrom SM, Denenberg VH, Palmiter RD (2005) Distinguishing whether dopamine regulates liking, wanting, and/or learning about rewards. *Behav Neurosci* 119(1):5–15
52. Goerendt IK, Lawrence AD, Brooks DJ (2004) Reward processing in health and Parkinson’s disease: neural organization and reorganization. *Cereb Cortex* 14(1):73–80
53. Ernst M, Kimes AS, London ED, Matochik JA, Eldreth D, Tata S et al (2003) Neural substrates of decision making in adults with attention deficit hyperactivity disorder. *Am J Psychiatry* 160(6):1061–1070
54. Blakemore SJ, Sirigu A (2003) Action prediction in the cerebellum and in the parietal lobe. *Exp Brain Res* 153(2):239–245
55. Tobler PN, O’Doherty JP, Dolan RJ, Schultz W (2006) Human neural learning depends on reward prediction errors in the blocking paradigm. *J Neurophysiol* 95(1):301–310
56. Gluck MA, Allen MT, Myers CE, Thompson RF (2001) Cerebellar substrates for error correction in motor conditioning. *Neurobiol Learn Mem* 76(3):314–341
57. Joel D, Niv Y, Ruppin E (2002) Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Netw* 15(4–6):535–547
58. Fiez JA, Petersen SE, Cheney MK, Raichle ME (1992) Impaired non-motor learning and error detection associated with cerebellar damage. A single case study. *Brain* 115(Pt 1):155–178
59. Cools R, Altamirano L, D’Esposito M (2006) Reversal learning in Parkinson’s disease depends on medication status and outcome valence. *Neuropsychologia* 44(10):1663–1673
60. Middleton FA, Strick PL (2001) Cerebellar projections to the prefrontal cortex of the primate. *J Neurosci* 21(2):700–712
61. Frank MJ, Claus ED (2006) Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev* 113(2):300–326
62. Ikai Y, Takada M, Mizuno N (1994) Single neurons in the ventral tegmental area that project to both the cerebral and cerebellar cortical areas by way of axon collaterals. *Neuroscience* 61(4):925–934