


# The Cerebellum Contributes to Prediction Error Coding in Reinforcement Learning in Humans

 Dana M. Huvermann,<sup>1,2</sup> Adam M. Berlijn,<sup>2,3</sup> Andreas Thieme,<sup>1</sup> Friedrich Erdlenbruch,<sup>1</sup> Stefan J. Groiss,<sup>4,5</sup> Andreas Deistung,<sup>6</sup> Manfred Mittelstaedt,<sup>2</sup> Elke Wondzinski,<sup>7</sup> Heike Sievers,<sup>7</sup> Benedikt Frank,<sup>1</sup> Sophia L. Göricke,<sup>1</sup> Michael Gliem,<sup>8</sup> Martin Köhrmann,<sup>1</sup> Mario Siebler,<sup>7,8</sup> Alfons Schnitzler,<sup>4,5</sup> Christian Bellebaum,<sup>2</sup> Martina Minnerop,<sup>3,4,5</sup> Dagmar Timmann,<sup>1\*</sup> and Jutta Peterburs<sup>2,9\*</sup>

<sup>1</sup>Department of Neurology and Center for Translational and Behavioral Neurosciences (C-TNBS), Essen University Hospital, University of Duisburg-Essen, Essen 45147, Germany, <sup>2</sup>Faculty of Mathematics and Natural Sciences, Heinrich Heine University Düsseldorf, Düsseldorf 40225, Germany, <sup>3</sup>Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich 52428, Germany, <sup>4</sup>Institute of Clinical Neuroscience and Medical Psychology, Medical Faculty & University Hospital Düsseldorf, Heinrich Heine University Düsseldorf, Düsseldorf 40225, Germany, <sup>5</sup>Department of Neurology, Center for Movement Disorders and Neuromodulation, Medical Faculty & University Hospital Düsseldorf, Heinrich-Heine University Düsseldorf, Düsseldorf 40225, Germany, <sup>6</sup>University Clinic and Outpatient Clinic for Radiology, Department for Radiation Medicine, University Hospital Halle (Saale), University Medicine Halle, Halle (Saale) 06120, Germany, <sup>7</sup>Department of Neurology and Neurorehabilitation, MediClin Fachklinik Rhein/Ruhr, Essen 45219, Germany, <sup>8</sup>Department of Neurology, Medical Faculty & University Hospital Düsseldorf, Heinrich Heine University Düsseldorf, Düsseldorf 40225, Germany, and <sup>9</sup>Institute for Systems Medicine & Department of Human Medicine, MSH Medical School Hamburg, Hamburg 20457, Germany

Recent rodent data suggest that the cerebellum—a region typically associated with processing sensory prediction errors (PEs)—also processes PEs in reinforcement learning (RL-PEs; i.e., learning from action outcomes). We tested whether cerebellar output is necessary for RL-PE processing in regions more traditionally associated with action-outcome processing, such as the striatum and anterior cingulate cortex. The feedback-related negativity (FRN) was measured as a proxy of cerebral RL-PE processing in a probabilistic feedback learning task using electroencephalography. Two complementary experiments were performed in humans. First, patients with chronic cerebellar stroke (20 male, 6 female) and matched healthy controls (19 male, 7 female) were tested. Second, single-pulse cerebellar transcranial magnetic stimulation (TMS) was applied in healthy participants (7 male, 17 female), thus implementing a virtual lesion approach. Consistent with previous studies, learning of action-outcome associations was intact with only minor changes in behavioral flexibility. Importantly, no significant RL-PE processing was observed in the FRN in patients with cerebellar stroke and in participants receiving cerebellar TMS. Findings in both experiments show that RL-PE processing in the forebrain depends on cerebellar output in humans, complementing and extending previous findings in rodents.

**Key words:** cerebellum; event-related potentials (ERPs); executive functions; lesion; noninvasive brain stimulation; performance monitoring

## Significance Statement

While processing of prediction errors in reinforcement learning (RL-PEs) is usually attributed to midbrain and forebrain, recent rodent studies have recorded RL-PE signals in the cerebellum. It is not yet clear whether these cerebellar RL-PE signals contribute to RL-PE processing in the forebrain/midbrain. In the current study, we could show that forebrain RL-PE coding is blunted when the cerebellum is affected across two complementary lesion models (patients with cerebellar stroke, cerebellar TMS). Our results support direct involvement of the cerebellum in RL-PE processing. We can further show that the cerebellum is necessary for RL-PE coding in the forebrain.

Received Oct. 17, 2024; revised March 12, 2025; accepted March 14, 2025.

Author contributions: D.M.H., A.M.B., S.J.G., C.B., Martina Minnerop, D.T., and J.P. designed research; D.M.H., A.M.B., A.T., F.E., and S.J.G. performed research; Manfred Mittelstaedt, E.W., H.S., B.F., M.G., M.K., M.S., and A.S. contributed unpublished reagents/analytic tools; D.M.H., A.D., and S.L.G. analyzed data; D.M.H., C.B., Martina Minnerop, D.T., and J.P. wrote the paper.

We thank the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG; project number: 437661157 awarded to J.P., D.T., and Martina Minnerop) for supporting this research. D.M.H. was partially supported by the Bernd Fink Foundation. We thank Greta Wippich for drawing the Experimental Procedure in

Figures 2A and 4A; Simone Lohbeck, Michael Klein, and Naomi Johanna Müller for their help with data acquisition; Thomas Ernst for his help with setting up the MRI protocol and conversion of MR data into BIDS format; and Beate Brol for her help with lesion symptom mapping.

\*D.T. and J.P. contributed equally to this work.

The authors declare no competing financial interests.

Correspondence should be addressed to Dana Huvermann at dana.huvermann@hhu.de.

<https://doi.org/10.1523/JNEUROSCI.1972-24.2025>

Copyright © 2025 the authors

## Introduction

In our fast-paced world, we need to constantly monitor our environment and our actions and choose according to the anticipated consequences of our actions. In such reinforcement learning contexts, we rely on external feedback (e.g., reward/success, punishment/failure) to acquire action-outcome associations and thereby improve our behavior. Thus, we need to learn to predict action outcomes, for which we rely heavily on processing prediction errors, i.e., the difference between predicted and actual outcomes.

Prediction errors in reinforcement learning contexts (RL-PEs) have mainly been linked to basal ganglia, midbrain, and prefrontal areas (Fouragnan et al., 2018). Predictive functions beyond the motor domain have also been proposed for the cerebellum (Ramnani, 2006; Sokolov et al., 2017). More recent studies have shown cerebellar activation patterns consistent with RL-PEs in both humans and rodents (Kostadinov and Häusser, 2022; Manto et al., 2024; Berlijn et al., 2024b). Moreover, psychiatric disorders with cerebellar involvement, such as schizophrenia, autism spectrum disorder, and major depression (Phillips et al., 2015), have been reliably associated with altered and/or impaired reinforcement learning (Balsters et al., 2017; Halahakoon et al., 2020; Katthagen et al., 2020). However, deficits cannot clearly be attributed to cerebellar dysfunction, as multiple brain areas are typically affected in these disorders. Causal evidence for cerebellar involvement in reinforcement learning in humans is scarce. Patients with cerebellar damage showed deficits in reversal, but not acquisition learning within reinforcement learning (Thoma et al., 2008; Nicholas et al., 2024). An initial study using electroencephalography (EEG) showed altered outcome/feedback processing in patients with cerebellar stroke without impaired acquisition learning (Rustemeier et al., 2016).

In humans, RL-PE processing is typically studied using feedback learning tasks. Here, participants have to learn through trial and error that, for example, one of several response options leads to a higher probability of monetary reward over punishment (Eppinger et al., 2008). EEG can be used to measure an approximation of activity in one of the main drivers of RL-PE processing, the anterior cingulate cortex (ACC; Fouragnan et al., 2018): the feedback-related negativity (FRN) is a frontocentral negative deflection in the event-related potential (ERP) that emerges ~250 ms after feedback onset (e.g., presentation of a reward or punishment; Miltner et al., 1997; San Martín, 2012). Reflecting the activity of dopaminergic target regions, such as the ACC and striatum (Holroyd and Coles, 2002; Hauser et al., 2014; Foti et al., 2015), FRN amplitudes covary with the estimated RL-PE at the single-trial level (Fischer and Ullsperger, 2013; Hoy et al., 2021; Rawls and Lamm, 2021).

Despite findings of RL-PE-like signals in the cerebellum in rodents and initial accounts of altered reinforcement learning in humans with cerebellar damage, acquisition learning seems to be largely unaffected in cerebellar lesion patients (Thoma et al., 2008; Rustemeier et al., 2016). In the present study, we therefore studied the interplay between the cerebellum and cerebral cortex with respect to RL-PE processing more directly by investigating the impact of cerebellar damage/disruption on cortical RL-PE coding in the FRN and learning success. This approach extends previous observations of RL-PE signals in the cerebellum toward the question whether these are necessary for intact cerebral RL-PE processing. In two experiments, we studied patients with chronic cerebellar stroke and used single-pulse transcranial magnetic stimulation (TMS) in healthy adults to create a “virtual

lesion.” Single-pulse TMS offers the advantage of examining effects of deficits with a high temporal precision. Reinforcement learning success was assessed using a probabilistic feedback learning task. RL-PE processing was assessed using the FRN on a trial-by-trial basis. If the cerebellum contributes to reinforcement learning, RL-PE processing as reflected in the FRN should be reduced in cerebellar stroke patients compared with healthy controls (Experiment 1) and for cerebellar single-pulse TMS (Experiment 2). Indeed, no significant RL-PE processing as indexed by the FRN was found in cerebellar stroke patients and for cerebellar TMS. Concerning behavior, only minor abnormalities in behavioral flexibility were observed, with reinforcement learning success generally preserved.

## Materials and Methods

### Experiment 1

#### Participants

Thirty-one adults with a chronic stroke restricted to the cerebellum were recruited from the university hospitals Essen and Düsseldorf as well as the Rhein-Ruhr Clinic in Essen-Kettwig, Germany. Only patients with a postacute stroke, i.e., at least 6 months after the stroke event (with one exception who was only 42 d poststroke), were included. Lesions had to be confined to the cerebellum. Thirty-three adults without stroke were recruited as controls. Inclusion criteria were no current psychiatric and no current or past neurological disease, no use of medications affecting the central nervous system, and no alcohol or illicit drug abuse. Five patients and four controls were excluded because they did not meet these inclusion criteria. In total, data from 26 patients (20 men and 6 women) and the 26 controls (19 men and 7 women) who provided the best match regarding demographic parameters entered the analyses. Three controls were excluded during matching to ensure that nontask-related parameters (Table 1) did not result in group differences in task-related variables. Means and standard deviations on the main demographic variables are listed in Table 1 (for details, see Table 2 for patients and Table 3 for controls). Of note, three patients with depression and antidepressant medication and three matched controls (who also had a clinical diagnosis of depression and antidepressant medication, roughly matched on BDI score) were included in the analyses, as we could not exclude this to be part of a cerebellar cognitive affective symptom (CCAS; Schmahmann and Sherman, 1998), and to ensure that target sample size was met.

Handedness was assessed with the Edinburgh Handedness Inventory (EHI; Table 1; Oldfield, 1971). According to LQ<sub>EHI</sub>, 19 patients and 21 controls were right-handed, 3 patients and 1 control were left-handed, and 4 patients and 4 controls were ambidextrous. IQ estimates were obtained using the Mehrfachwahl-Wortschatz-Test-B score (MWT-B; multiple choice vocabulary test; Table 1; Merz et al., 1975). As depression might affect feedback processing (Keren et al., 2018) and has a higher incidence poststroke (Robinson and Jorge, 2016), we assessed depression using the Beck Depression Inventory II (BDI; Beck et al., 1996).

Figure 1 shows the overlaid lesion regions for all 26 patients. Overall, 21 patients had a stroke in the posterior inferior cerebellar artery (PICA) territory (9 left, 9 right, 3 bilateral), 3 patients had a stroke in the superior

**Table 1. Demographic data for patients and controls as well as group comparisons**

Variable	$M_{\text{controls}}$ (SD)	$M_{\text{patients}}$ (SD)	$t$	df	$p$
Age	56.4 (12.7)	56.2 (12.1)	0.04	49.86	0.964
Education (years)	13.5 (2.3)	14.1 (2.9)	0.90	47.44	0.371
LQ <sub>EHI</sub>	70.9 (47.2)	56.6 (54.0)	1.01	49.11	0.316
MWT-B (IQ)	118.8 (16.1)	113.4 (13.9)	1.28	48.91	0.205
WML rating	0.31 (0.47)	0.69 (0.79)	2.14	40.81	0.039
BDI-II	5.3 (7.5)	6.3 (9.1)	0.42	48.17	0.679

$n_{\text{controls}} = 26$ ,  $n_{\text{patients}} = 26$ .  $M$ , mean; SD, standard deviation. Age and education are given in years. LQ, laterality quotient; EHI, Edinburgh Handedness Inventory; MWT-B, Mehrfachwahl-Wortschatz-test B (a vocabulary-based German intelligence test); WML, white matter lesion; BDI-II, Beck depression inventory II.

**Table 2. Patients' and lesion characteristics**

ID	Age	Sex	Edu.	Hnd.	Vascular territory	Time since stroke (years)	Affected cerebellar regions	Additional information
005	55–59	m	16–20	r	PICA-L (lacunar)	0–5	Left Crus II, left Crus I, left VIIIa	
011	45–49	m	16–20	r	PICA-R	0–5	Right Crus II, right VIIIa, right VIIb, right VIIIb, right IX, right Crus I	External MRI
012	30–34	m	16–20	r	PICA-R	0–5	Right VIIIa, right VIIb, right VIIIb, right IX, right Crus II	
014	65–69	m	11–15	a	PICA-R	21–25	Right VIIb, right Crus II, right VIIIa, right VIIIb, right IX, vermal VIIIa, vermal VIIIb	
018	55–59	m	11–15	a	PICA-L	11–15	Left Crus I, left Crus II, left VIIIa, left VIIb, left VIIIb, left VI	
020	70–74	m	11–15	r	SCA-R	11–15	Right I–IV, right V, right Dentate	
023	55–59	f	11–15	l	PICA-R	16–20	Right Crus I, right VIIIb, right IX, right Crus II, right VIIb, right VIIIa	
024	60–64	m	11–15	r	PICA-L+R (lacunar)	6–10	Right VIIb, left VIIb	
026	65–69	f	6–10	r	PICA-R, PICA-L (lacunar)	6–10	Right VIIIa, right VIIb, right IX, right Crus II, right VIIIb, left VIIIa, left IX, left VIIIb	Migraine
029	50–54	m	16–20	a	SCA-R	11–15	Right VI, right V, right Dentate, right Crus I	Recent history of AD, negative BAI
031	45–49	m	11–15	r	PICA-R (lacunar)	0–5	right Crus I, right Crus II, right VIIIa	
033	50–54	f	11–15	r	PICA-L, SCA-R (lacunar)	6–10	Left Crus I, left Crus II, left VIIb, right Crus I, left VIIIa, vermal Crus II	MDD, intake of antidepressants
035	60–64	m	16–20	a	PICA-L	16–20	Left VIIIa, left VIIb, left VIIIb, left Crus II, left IX, left Crus I	MDD, antidepressants
037	55–59	f	16–20	r	PICA-L	0–5	Left IX, left VIIIb, left VIIb, left VIIIa, left Crus II, left Dentate	External MRI
038	55–59	f	11–15	r	PICA-L, PICA-R (lacunar)	6–10	Left Crus II, left VIIb, left VIIIa, left IX, left Crus I, left VIIIb	
039	60–64	m	16–20	l	PICA-R (lacunar)	n/a <sup>a</sup>	Right Crus I	External MRI
040	60–64	m	16–20	r	PICA-L	0–5	Left Crus II, left VIIb, left VIIIa, left VIIIb, left Crus I, left IX	Not a native German speaker
042	50–54	m	11–15	r	PICA-R	6–10	Right Crus II, right VIIb, right Crus I, right VIIIa	
046	18–24	m	11–15	r	PICA-R	0–5	Right VIIb, right VIIIa, right VIIIb	External MRI
047	65–69	m	16–20	r	PICA-R	0–5	Right Crus II, right Crus I, right VIIb, right VIIIa, right VIIIb, right IX	
048	65–69	m	16–20	l	PICA-L	6–10	Left Crus II, left VIIb, left VIIIa, left VIIIb, left Crus I, left IX, left dentate	
055	70–74	m	6–10	r	PICA-L	6–10	Left Crus II, left Crus I	
056	65–69	m	11–15	r	SCA-L	6–10	Left V, left VI, left I–IV, left dentate	
058	45–49	f	11–15	r	PICA-L (lacunar)	0–5	Left VIIIb, left IX	MDD, childhood diagnosis of ADD, antidepressants
060	50–54	m	11–15	r	SCA-L, PICA-R	16–20	Left V, left VI, right Crus II, left I–IV, right VIIb, left Dentate	
061	55–59	m	16–20	r	PICA-L	0–5	Left Crus II, left Crus I, left VIIb, left VIIIa, left VIIIb, left IX	

Age is given in years. m, male; f, female; edu., years of education; hnd., handedness according to Edinburgh Handedness Inventory; l, left; r, right; a, ambidextrous; PICA, posterior inferior cerebellar artery; SCA, superior inferior artery; SARA, scale for the assessment and rating of ataxia (maximum score = 40); MDD, major depressive disorder; AD, anxiety disorder; ADD, attention deficit disorder; BAI, Beck anxiety inventory (Beck et al., 1988). For affected regions, only those that made out >1% of total lesion volume were included. Regions are sorted according to percentage of total lesion volume. Age, education, and time since stroke are given in ranges to comply with data protection requirements. <sup>a</sup>Time since stroke was unknown in this case, as the lacunar stroke was an incidental finding.

cerebellar artery (SCA) territory (1 left, 2 right), and 2 patients had a stroke in both the PICA and SCA territory (1 in the left PICA and right SCA territory, 1 in the right PICA and left SCA territory). Lesions were thus unilateral in all but five patients (Table 2). Images of individual lesions are provided in Extended Data Figure 1–1. Mean time between cerebellar infarct and participation in the experiment was 8.4 years (SD = 6.0 years, range from 1.5 months to 22 years; unknown in one case).

Both patients and controls were assessed for clinical neurological symptoms. While patients showed overall higher scores on the scale for the assessment and rating of ataxia (SARA; Schmitz-Hübsch et al., 2006) than controls ( $t_{(47.01)} = 2.74$ ,  $p = 0.009$ ), no participants showed deficits in oculomotor function in the neurological examination (which might have affected task performance).

All participants gave written informed consent prior to participation. They received monetary compensation for participation and were reimbursed for travel costs. The experiment was preregistered on the Open Science Framework (OSF; <https://osf.io/rd3xb>), conducted in accordance with the ethical principles for medical research involving human subjects outlined in the Declaration of Helsinki, and approved by the Ethics Committees at the Faculty of Medicine of Heinrich-Heine-University Düsseldorf and at the University Hospital Essen.

#### Procedure

The experiment usually took place on 2 consecutive days. The temporal gap between sessions was longer for three patients (56, 29, and 43 d) and

two controls (9 and 2 d) due to participants' time constraints or scheduling issues. The experimental task with EEG was conducted on both days with different versions (short and long feedback delay version, respectively; see below for details). On the first day, we initially obtained informed consent and participants filled in a demographic questionnaire, the EHI, the BDI-II, and the MWT-B. Following EEG preparations, participants were informed about EEG artifacts and how to avoid them.

Subsequently, participants completed one of two feedback delay versions of a probabilistic feedback learning task as described by Eppinger et al. (2008; see Fig. 2A for the experimental procedure), which was conducted using Presentation software (version 20.0, Neurobehavioral Systems). Order of the versions was counterbalanced among participants. Figure 2B illustrates the sequence and time course of stimulus presentation in one trial. The task consisted of 8 blocks of 40 trials, thus 320 trials in total. Five practice trials with different stimuli were provided. Each trial began with a fixation cross presented for 500–1,500 ms. Next, one of four abstract stimuli (Chinese characters and radicals) was presented for 1,500 ms. Participants had to respond by pressing the left or right button on a response box (Cedrus RB-740, Cedrus Corporation) within a response time window of 3,000 ms. Choices were highlighted on the screen for 200 ms, followed by a black screen for 500 ms in the short delay condition and 6,500 ms in the long delay condition. Different delay durations were used as previous studies had shown differential involvement of cerebral brain areas depending on feedback delay (Foerde and Shohamy, 2011). While immediate delay

**Table 3. Characteristics of controls for Experiment 1 (patient study)**

ID	Age	Sex	Handedness (EHI)	Years of education	Additional information
001	55–59	f	Right	11–15	
002	75–79	m	Right	16–20	Excluded (matching)
004	60–64	m	Right	11–15	Excluded (developmental venous anomaly)
006	70–74	m	Right	11–15	
007	55–59	f	Right	11–15	
008	55–59	m	Right	6–10	MD, intake of antidepressants
009	70–74	m	Right	11–15	
010	50–54	f	Right	11–15	
013	55–59	m	Right	16–20	
017	70–74	f	Right	11–15	
019	60–64	m	Right	11–15	
021	50–54	m	Right	16–20	
022	55–59	f	Right	16–20	MD, intake of antidepressants
025	55–59	m	Right	11–15	Excluded (moderate brain volume loss)
027	70–74	m	Right	16–20	
028	45–49	f	Right	11–15	MD, intake of antidepressants
030	45–49	m	Right	11–15	
032	55–59	m	Right	16–20	Excluded (matching)
034	55–59	m	Right	11–15	
036	25–29	m	Right	6–10	Excluded (intake of antidepressants)
043	18–24	m	Right	11–15	
044	55–59	m	Left	11–15	
045	65–69	m	Ambidextrous	11–15	
049	70–74	m	Right	11–15	Not a native German speaker
050	50–54	f	Right	6–10	
051	50–54	m	Right	11–15	MD, intake of antidepressants, excluded (matching)
052	18–24	m	Ambidextrous	6–10	
054	55–59	m	Ambidextrous	11–15	
057	70–74	m	Right	11–15	Excluded (extensive white matter lesions)
059	55–59	m	Ambidextrous	16–20	
062	50–54	m	Right	11–15	
063	55–59	m	Right	11–15	
064	65–69	m	Right	11–15	

Age is given in years. m, male; f, female; EHI, Edinburgh Handedness Inventory; SARA, Scale for the Assessment and Rating of Ataxia; MD, major depression. Age and education are given in ranges to comply with data protection requirements.

activated areas typically associated with reward processing, such as the striatum, activations for delayed feedback shifted toward the hippocampus (Foerde and Shohamy, 2011). This shift was also apparent in FRN, with a decreased FRN amplitude with longer feedback delays (Peterburs et al., 2016). This shift in activation could affect potential deficits in cerebellar patients due to differences in connectivity of these brain regions with the cerebellum. Finally, feedback was displayed for 1,000 ms. Feedback consisted of either the display of “+20 ct” in green font as positive feedback or “–10 ct” in red font as negative feedback.

Two of the four stimuli were linked to random feedback (50% positive and 50% negative regardless of response), while the other two were linked to contingent feedback. Here, correct responses were followed by positive feedback 90% of the time and by negative feedback 10% of the time (and vice versa for incorrect responses). Correctness was balanced for the two response buttons, so that for one of these stimuli, the chance of positive feedback was higher for the left button, while for the other stimulus, the chance of positive feedback was higher for the right button. In case a participant exceeded the learning criterion of 65% correct answers by the second of eight blocks, a new stimulus set was provided to increase the number of prelearning trials. This was the case for eight patients and eight controls in one feedback delay condition/session (of which 1 and 5,

respectively, were second sessions) and for six patients and eight controls in both conditions. In case a participant did not exceed the learning criterion until the eighth and last block, a ninth block was added to increase the number of postlearning trials. This was the case for three patients and two controls in one of the two conditions.

Following this task on Day 1, participants underwent cranial MRI and a clinical neurological examination.

On the second day, following EEG preparations, the remaining other version (short or long feedback delay version) of the probabilistic feedback task was completed. Versions were counterbalanced between sessions. Two different stimulus sets were used per session, and order was counterbalanced. Responses (choice, choice accuracy) and response times were recorded during the experiment.

#### EEG recording and preprocessing

EEG was recorded from 28 active Ag/AgCl electrodes (F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, PO9, O1, Oz, O2, PO10) positioned on a BrainCap (Brain Products) according to the 10–20 system. FCz was used as an on-line reference, and AFz was used as ground electrode. Fp1 was used as vertical electrooculogram (vEOG) and an electrode was placed next to the outer canthus of the left eye as horizontal electrooculogram (hEOG). Impedances were kept below 25 k $\Omega$ . Data were amplified with a BrainAmp DC amplifier and recorded at 1,000 Hz using BrainVision Recorder 1.21 (Brain Products). Data preprocessing was performed using BrainVision Analyzer 2 software (version 2.2, Brain Products) and MATLAB (MathWorks).

First, data were rereferenced to the mastoid electrodes and FCz was reestablished. Next, a DC detrend was applied and data were filtered using zero phase shift Butterworth filters with a low cutoff of 1 Hz and high cutoff of 30 Hz, as well as a notch filter at 50 Hz to remove powerline artifacts. Subsequently, we removed vertical and horizontal eye movement artifacts using a semiautomatic Ocular Correction ICA as implemented in BrainVision Analyzer 2. vEOG was used for blinks and vertical activity, and hEOG for horizontal activity. The first 177.2 s were used for ICA. We then segmented data from the start marker of the experiment to end of experiment, and segmented them around the feedback markers, starting 200 ms before and ending 600 ms after each marker. Only feedback markers for learnable stimuli were segmented. Baseline correction was performed based on the 200 ms preceding feedback onset, followed by automated artifact rejection. Segments with a voltage step exceeding 50  $\mu$ V/ms, an amplitude above 100  $\mu$ V or below –100  $\mu$ V, or activity not exceeding 0.1  $\mu$ V were excluded. Single-trial data were then exported via generic data export. On average, 1.1% of segments (SD = 2.5%) were rejected. Additionally, data for learnable stimuli were averaged and exported according to feedback valence (positive, negative) and feedback delay (short, long).

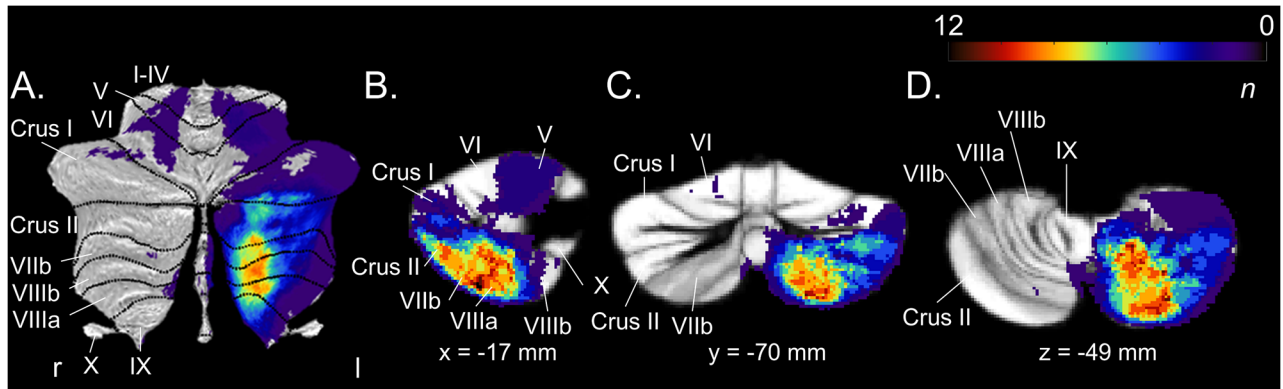
The EEG system had to be switched from an actiCAP system to a newer actiCAP snap system after the first 27 participants due to a defect in impedance measurement.

In MATLAB, peak detection was performed on the averaged data separately for each condition [feedback valence (positive, negative)  $\times$  feedback delay (short, long)]. The FRN was defined as the local maximal negative peak within the time window between 200 and 350 ms at electrode site FCz (Sambrook and Goslin, 2015). If no local maximum/minimum could be detected, the corresponding single-trial segments were excluded. For the single-trial data, the mean amplitude in a time window of 40 ms around the respective FRN latency determined by the peak detection on the averaged data was extracted (Meadows et al., 2016).

#### Prediction error estimation

Prediction errors on each trial were estimated based on choices participants made and feedback they received using a reinforcement learning model (Sutton and Barto, 2018) consistent with previous studies (Ichikawa et al., 2010; Fischer and Ullsperger, 2013; McDougle et al., 2019) which has been shown to be highly correlated with the gold standard (i.e., subjective ratings; Ichikawa et al., 2010). We modeled action values  $Q$  and PEs  $\delta$  based on the actually received feedback  $R$  and



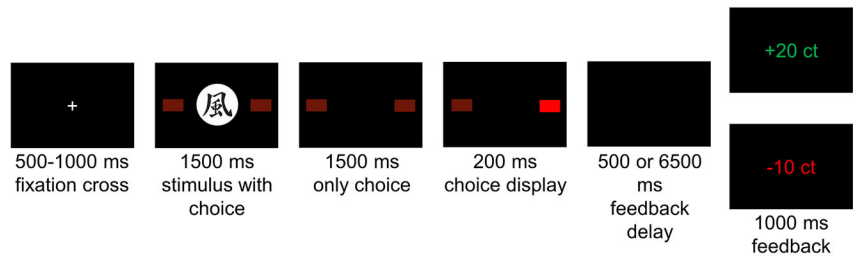


**Figure 1.** Overlap plot of all lesions in the patient group ( $n = 26$ ) superimposed on (A) a cerebellar flatmap (Diedrichsen and Zotow, 2015) and in 2D (B) sagittal, (C) coronal, and (D) axial views. Lesions on the right side were mirrored to the left side. Color code shown on the top right denotes total lesion overlap (from purple = 0 to red = 12). Individual lesions are depicted in Extended Data Figure 1-1.

## A. Experimental setup



## B. Time course and sequence of stimulus presentation



**Figure 2.** Experimental procedure of Experiment 1 (patient study). **A**, Experimental setup. EEG was recorded while participants performed the probabilistic feedback task. **B**, Time course and sequence of stimulus presentation in one trial of the feedback learning task. After a fixation cross was presented for 500–1,000 ms, one of four stimuli was presented, toward which participants were required to respond by pressing the left or right button on a response pad within 3,000 ms. The stimulus was only shown for the first 1,500 ms. After the response, the respective choice was highlighted on screen for 200 ms, followed by either 500 or 6,500 ms of blank screen (depending on feedback delay version). Positive (“+20 ct”) or negative feedback (“−10 ct”) was then presented on screen for 1,000 ms. Participants needed to learn by trial and error whether one of the choices was related to higher chance of positive/negative feedback depending on stimulus. Feedback for two of the stimuli had a 90% contingency, while for the other two it had a 50% (random) contingency. A total of 320 trials were used in the task.

participants’ chosen response  $a$ , using a Rescorla–Wagner model (Rescorla and Wagner, 1972; Wagner and Rescorla, 1972):

$$Q_{a,t+1} = Q_{a,t} + \alpha * \delta_t,$$

$$\delta_t = R_{a,t} - Q_{a,t}.$$

To model response probabilities, we used a softmax function (Sutton and Barto, 2018), which estimates the probability  $p$  of the chosen action with the estimated action values  $Q$  per action  $a$  and time point  $t$  (in this case, the trial):

$$p_{a,t} = \frac{e^{\beta * Q_{a1,t}}}{e^{\beta * Q_{a1,t}} + e^{\beta * Q_{a2,t}}}.$$

The function *fmincon* provided by MATLAB was used to fit this model to the data via minimizing the negative sum of log-likelihoods minus a gamma distribution of  $\beta$  with a shape parameter of 2 and scale parameter of 3 (as to penalize high  $\beta$ ; McDougle et al., 2019). We estimated a learning rate  $\alpha$  as well as an inverse temperature  $\beta$  for exploration behavior, separately for each stimulus and reward and punishment. We allowed  $\alpha$  to assume any value between 0 and 1 and  $\beta$  to assume any value between 0 and 50.

### Experimental design and statistical analysis

This study was preregistered to OSF (<https://osf.io/rd3xb>). Necessary sample size was determined via power analysis to be  $N = 48$ , i.e., 24 per

group, as detailed in the preregistration. Required sample size was thus matched ( $n = 26$  per group). Raw data and code used for preprocessing and analysis are available from <https://osf.io/cqf97>.

Analysis focused on differences in the FRN between controls and patients (on group level), especially in relation to coding of RL-PEs. As the signed RL-PE overlaps with feedback valence, we split the signed RL-PE into the unsigned RL-PE (on a scale from 0/low to 1/high) and feedback valence (positive, negative), which were used as separate predictors in the analysis. We further examined learning success as reflected in choice accuracy and choice switching as an index of behavioral flexibility. The analysis was restricted to stimuli with a 90% contingency, as participants were not able to learn in the 50% contingency condition.

Data were analyzed in R (version 4.2.3; R Core Team, 2023) using RStudio (version 2023.3.0.386, Posit Team, 2023). Concerning choice accuracy, the preregistered ANOVA was performed. For the FRN, since only 14 patients and 12 controls exceeded the learning criterion of >65% correct responses within at least one block of the task in either version, the preregistered ANOVA analysis with learning phase (pre-/postlearning) as a factor was not feasible. With the factor learning phase, we had aimed to investigate to what extent feedback processing changed over the course of the task as participants learned which responses resulted in a higher chance of reward/punishment. Instead, we decided to pursue a single-trial-based analysis approach using LME models including the trial-by-trial unsigned RL-PE. Analyses based on single trials have increasingly been used in recent studies as they offer the possibility to use variables that vary from trial to trial as factors in the statistical analysis (Volpert-Esmond et al., 2021). LME analyses based on single-trial data have also been shown to deliver less biased results compared

with ANOVAs based on averaged data (Heise et al., 2022). The lme4 library (version 1.1-32; Bates et al., 2015) was used for LME modeling, and the lmerTest library (version 3.1-3; Kuznetsova et al., 2017) was used to evaluate statistical significance. Significance was evaluated using restricted maximum likelihood with  $p$  values computed using Satterthwaite approximation, following the findings by Luke (Luke, 2017). While we initially tested a maximal fit for random effects, in case of singular fit, we reduced the originally maximal random effects structure up to the random intercept and highest-order interaction as random slope per participant (Brauer and Curtin, 2018). For all LME analyses, outliers were identified via Cook's distance (Cook, 1977) using the influenceME package (version 0.9-9; Nieuwenhuis et al., 2012) and an outlier criterion of  $4/(n-p-1)$ , where  $n$  is the number of subjects and  $p$  is the number of fixed effects. Significant interactions were followed up using simple slope analyses via the interactions library (version 1.1.5; Long, 2019).  $p$  values were Bonferroni-corrected according to the number of simple slopes in the respective analysis.

**Choice accuracy.** We conducted a mixed ANOVA with the factors group (patients, controls), feedback delay (short, long), and block (1–8). Significant effects were followed up with Bonferroni-corrected  $t$  tests using the *emmeans\_test* function (Lenth, 2025). No participant exceeded the outlier criterion of  $M \pm 2.5$  SD per feedback delay/study session.

**Choice switching.** As an additional behavioral measure, we analyzed whether choice switching following feedback was influenced by the categorical fixed effects feedback valence (−0.5: negative, 0.5: positive), response type (−0.5: false, 0.5: correct), group (−0.5: control, 0.5: patient), feedback delay (−0.5: short, 0.5: long), and the continuous effect block which was scaled via the built-in *scale* function. Choice switching for a given trial was defined as whether the choice for the current stimulus was switched (choice switching = 1) or sustained (choice switching = 0) in the next trial that the same stimulus was presented in. The variable was scaled via the built-in *scale* function. We also included all interactions of these factors as fixed effects. No participants exceeded our Cook's distance criterion. The model equation was as follows:

$$\text{choice switching} \sim 1 + \text{feedback valence} * \text{response type} * \text{group} \\ + \text{feedback delay} * \text{block} + (1 + \text{feedback valence}: \\ \text{response type}:\text{feedback delay}:\text{block}|\text{subject})$$

**FRN.** For FRN amplitudes, we again employed an LME model with the fixed effects feedback valence (negative: −0.5, positive: 0.5), group (−0.5: control, 0.5: patient), feedback delay (−0.5: short, 0.5: long), and the continuous fixed effect unsigned RL-PE which was the absolute of the signed RL-PE minus 0.5 (thus with minimal values of −0.5 and maximal values of 0.5). We also included all interactions of these factors as fixed effects. Here, we deviated from the preregistration (which only included the signed RL-PE), because analyzing the signed RL-PE in an LME model is confounded by valence effects and disregards possible U-shaped relations which are identifiable by separating feedback valence and unsigned RL-PE (i.e., RL-PE magnitude). Initial convergence issues were solved via changing the optimizer to *bobyqa*. Four controls were excluded due to exceeding the Cook's distance criterion. The model equation was as follows:

$$\text{FRN} \sim 1 + \text{unsigned PE} * \text{feedback valence} * \text{group} * \text{feedback delay} \\ + (1 + \text{unsigned PE}:\text{feedback valence}:\text{feedback delay} + \\ \text{unsigned PE}:\text{feedback valence} + \text{unsigned PE}:\text{feedback delay} \\ + \text{feedback valence}:\text{feedback delay}|\text{subject})$$

#### Structural MRI and lesion symptom mapping

For 21 patients and all controls, a 3D T1-weighted magnetization-prepared rapid acquisition gradient-echo (MPRAGE) sequence was acquired [176 slices, repetition time (TR), 2,530 ms; echo time (TE), 2.27 ms; inversion time (TI), 1,100 ms; flip angle (FA), 7°; voxel size,  $1 \times 1 \times 1$  mm; acceleration factor, 2 (GRAPPA); field of view,  $256 \times$

$256$  mm; acquisition time (TA), 6:03 min:s]. A MAGNETOM Vida 3T system (Siemens Healthcare) with a 64-channel coil was used. For the remaining five patients, an MR scan was not possible due to implants ( $n=4$ ) or claustrophobia ( $n=1$ ), and instead, existing diagnostic structural MR images were used.

We first confirmed that lesions were isolated to the cerebellum, which was also reconfirmed by an experienced neuroradiologist (SGö). T2-weighted images were also assessed for white matter lesions (see below).

Non-normalized 3D T1 images were first manually aligned to the AC-PC line. Cerebellar, postischemic lesions were then manually traced and saved as regions of interest using MRICron (<https://www.nitrc.org/projects/mricron>). Next, the cerebellum was isolated, and datasets were segmented using the *suit\_isolate\_seg* function provided by SUIT toolbox (<https://www.diedrichsenlab.org/imaging/suit.htm>). Isolation masks were manually corrected. Datasets were then normalized with the function *suit\_isolate\_mask*, using the lesion mask as optional input, thus ignoring the respective area(s). Finally, lesion ROIs were transformed via *suit\_reslice* into the spatially unbiased atlas template of the cerebellum (SUIT; Diedrichsen, 2006).

For statistical analysis of whether deficits corresponded to specific lesion locations, voxel-based lesion symptom mapping (vLSM) was conducted using NPM as implemented with MRICron (Stoodley et al., 2016; Timmann et al., 2022). For this purpose, all lesion ROIs on the right side were mirrored to the left side (in the five patients with bilateral lesions, the side with the larger lesion was considered). For one subject (sub-060) with bilateral lesions, the lesion of higher interest for our cognitive task in posterolateral regions (Crus II, lobule VIIb) was mirrored to the left side instead of the larger lesion in anterior cerebellar motor regions. vLSM compares for each voxel, whether patients with this voxel affected differ from patients with this voxel unaffected within a variable of interest. A Brunner–Munzel test was employed (Brunner and Munzel, 2000) with a statistical threshold of  $p < 0.05$ . Only variables with statistical differences between patients and controls were considered, which in our case was the interaction effect between RL-PE, feedback valence, and group onto FRN amplitude. We collapsed this effect into one variable per participant by taking the difference between the mean FRN amplitude for high RL-PE ( $\geq -0.5$ ) and low RL-PE ( $< 0.5$ ) for negative feedback (because FRN amplitudes differed between high and low RL-PE only for negative feedback valence in controls). Signs were reversed for the analysis, as lower values are related to more dysfunction in vLSM (while in our case positive values were associated with more dysfunction). Clusters of voxels with significant effects were extracted using MRICroGL (Brett et al., 2001), considering clusters larger than  $32$  mm<sup>3</sup>. Affected lobules and nuclei were defined based on the probabilistic atlases of the human cerebellum by Diedrichsen et al. (2009, 2011).

#### White matter lesion assessment

3D dark fluid T2-weighted spin-echo sequences [SPACE; 160 slices; TR, 7,000 ms; TE, 428 ms; TI, 2,050 ms; voxel size,  $1 \times 1 \times 1$  mm; acceleration factor, 2 (GRAPPA), field of view,  $256 \times 256$  mm; TA, 5:24 min:s] were acquired and examined for white matter lesions. They were rated following Wahlund et al. (2001; Table 1). Even though we excluded patients with particularly pronounced and widespread white matter lesions, we still found higher white matter lesion ratings for cerebellar stroke patients than controls. White matter lesions have previously been shown to contribute to deficits in cognition (Filley and Fields, 2016), although findings specifically concerning reinforcement learning appear as yet lacking. To exclude effects of WMLs in our data, we checked our cognitive scores (CCAS) and did not find any general cognitive deficits in patients compared with controls ( $t_{(40,89)} = 1.03$ ,  $p = 0.310$ ). Our finding concerning RL-PE processing in stroke patients additionally coincides with our findings for cerebellar TMS in Experiment 2, where cerebellar and control stimulation was applied within-subject, excluding between-subject factors like white matter lesions as a cause. It thus seems likely that the deficits in RL-PE processing in our patients are caused by the cerebellar stroke itself.

#### Quantitative susceptibility mapping

In an additional analysis, we examined whether lesions within the dentate nucleus had a special impact on RL-PE processing, as the dentate

nucleus constitutes the main output of the cerebellum, with more than half of its projection relating to non-motor functions (Palesi et al., 2021). To identify the dentate nucleus based on its high iron levels (Deistung et al., 2016), quantitative susceptibility mapping (QSM) was conducted based on data collected using a multi-echo gradient-echo scan [176 axial slices; TR, 27 ms; TE<sub>1-4</sub>, 3.66 ms/9.74 ms/15.83 ms/21.91 ms; FA, 15°; voxel size, 0.9 × 0.9 × 0.9 mm; acceleration factor, 2 (GRAPPA); field of view, 230 mm × 230 mm; TA, 8:15 min:s] as described in Deistung et al. (2022). Outlines of the dentate nucleus as well as dentate nucleus lesion (if present) were manually drawn in ITK-SNAP (Yushkevich et al., 2006). We identified six patients who had a lesion in the dentate nucleus, with two of those previously classified as impaired concerning the differentiation between negative low and high RL-PE in FRN. We coregistered both dentate nucleus regions of interest and lesions to the T1 images and then normalized and resliced them into SUIT space. Functions included within the SUIT toolbox were used, i.e., *suit\_normalize\_dentate* for normalization and *suit\_reslice\_dartel* for reslicing. However, as overlaps between lesions were too few and did not allow meaningful statistical analysis, we abstained from an additional analysis and instead provide images of the individual dentate lesions in Figure 3.

## Experiment 2

### Participants

Twenty-nine healthy adults were recruited for participation. Four participants completed only one of two sessions and were thus excluded; one participant was excluded due to stimulation at a false output strength in one of the sessions. Thus, data from 24 participants (7 men, 17 women) with a mean age of 23.3 years (SD = 2.9 years, range from 19 to 30 years) were analyzed. Handedness was assessed with the EHI (Oldfield, 1971), with a mean LQ score of 62.3 (SD = 53.3, range from -85.7 to 100.0). According to LQ<sub>EHI</sub>, 20 participants were right-handed, 2 left-handed, and 2 ambidextrous. All participants reported no neurological or psychiatric diseases and no metal implants in or near their head. Further exclusion criteria were pregnancy, alcohol or illicit substance abuse, and intake of psychotropic medication. IQ estimates were obtained using the MWT-B (Merz et al., 1975), yielding a mean IQ of 103.5 (SD = 15.4). Participants received monetary compensation for participation in two sessions.

All participants gave written informed consent prior to participation. The experiment was conducted in accordance with the ethical principles for medical research involving human subjects outlined in the Declaration of Helsinki and approved by the Ethics Committee at the Faculty of Medicine of Heinrich-Heine-University Düsseldorf.

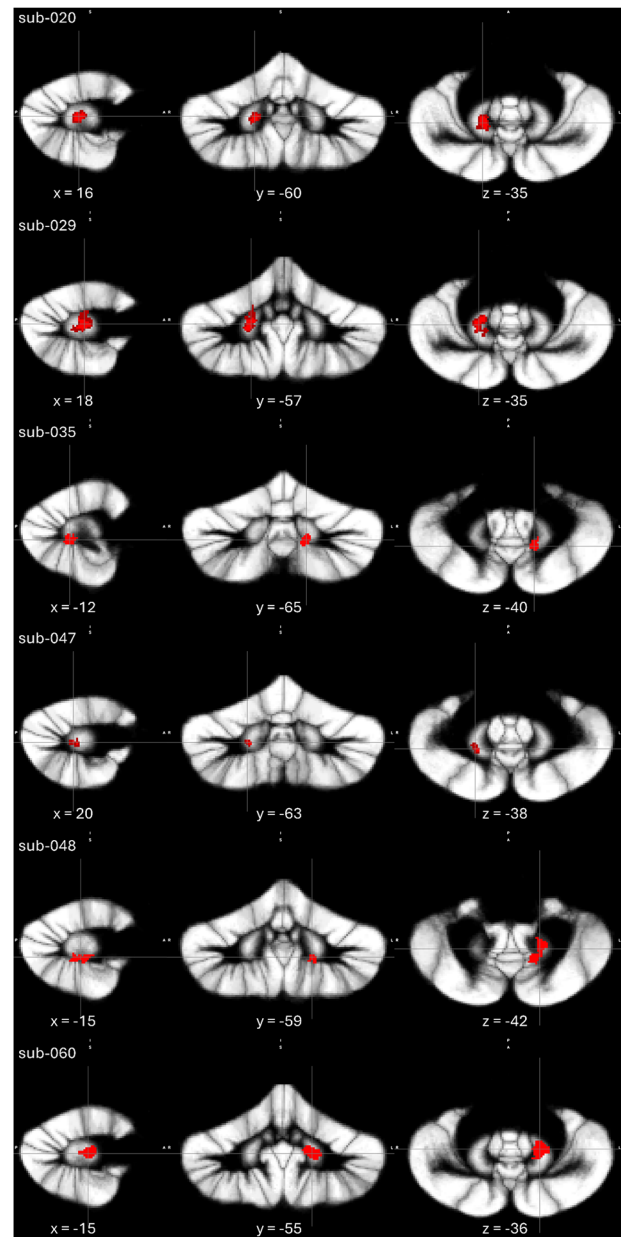
### Procedure

The experiment took place on 2 separate days with at least 48 h in between to decrease repetition effects ( $M = 101.6$  d,  $SD = 152.1$  d, range from 2 to 448 d). Note that due to a technical defect of the TMS system and a consequent pause of experiments, the time between sessions was exceptionally long for five individuals included in the analysis (362–448 d). Without these participants, the average time between sessions was 26.6 d ( $SD = 28.0$  d, range from 2 to 98 d).

While one session of Experiment 2 comprised the experimental task with vertex (control) stimulation, the other session comprised the cerebellar stimulation. Order for stimulation site was counterbalanced.

After participants arrived in the lab, informed consent was obtained and they filled in a demographic questionnaire, the EHI, and the MWT-B. Following EEG and EMG preparations and the subsequent motor threshold estimation, we placed the double cone TMS coil on their head with a custom mounting and further secured it with an elastic band (see below for a detailed description; Fig. 4A). Before and after the experimental task, an additional Flanker task was performed for which results are reported elsewhere (Berlijn et al., 2024a).

Participants completed a probabilistic feedback learning task that closely followed procedures as described for Experiment 1. Figure 4B illustrates the sequence and time course of stimulus presentation in each trial. The task consisted of six blocks of 56 trials, thus 336 trials

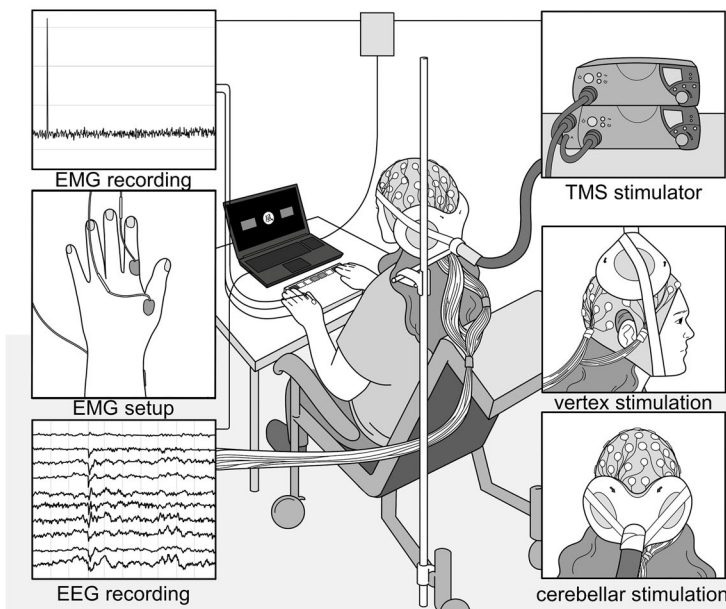


**Figure 3.** Individual, unmirrored dentate nucleus lesions for Experiment 1 (patient study), normalized to SUIT space, are presented in sagittal (left column), coronal (middle column), and axial view (right column) for patients with dentate lesions (sub-004, sub-005, sub-019, sub-022, sub-026, and sub-036). Lesion are marked in red.

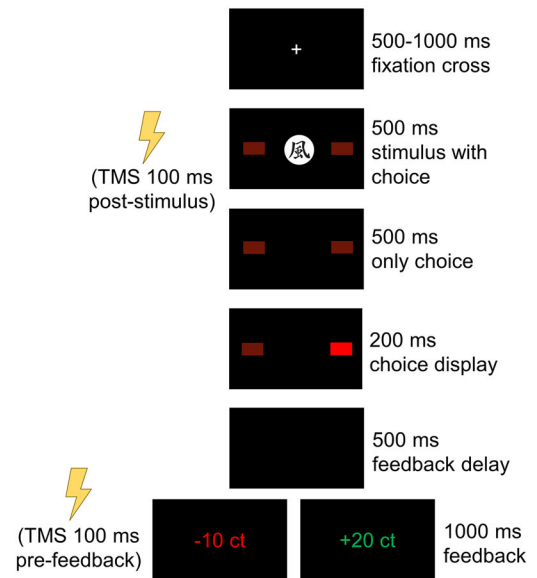
in total. Again, five practice trials with different stimuli were provided. Due to the younger sample, stimulus presentation was reduced to 500 ms and the response time window was shortened to 1,000 ms. Only short feedback delays (i.e., 500 ms) were used. Two stimuli were again linked to random feedback while the other two stimuli were linked to contingent feedback. For the contingent stimuli, correct responses were followed by positive feedback in 80% of the cases and by negative feedback in 20% of the cases (vice versa for incorrect responses). In both contingency conditions, TMS was delivered 100 ms poststimulus for one stimulus and 100 ms prefeedback for the other. In case a participant had learnt so fast that they exceeded the learning criterion of 75% correct answers by the second of six blocks, a new stimulus set was provided to increase the number of prelearning trials. This was the case for seven participants in one condition (of which four were second sessions) and for one participant in both conditions. In case a participant did not exceed the learning criterion until the sixth and last block, a seventh



## A. Experimental setup



## B. Time course and sequence of stimulus presentation



**Figure 4.** Experimental procedure of Experiment 2 (TMS study). **A**, Experimental setup. A double cone coil was placed on either the left cerebellum (1 cm down and 3 cm to the left of theinion) or vertex depending on session. Simultaneously, EEG and EMG were recorded. **B**, Time course and sequence of stimulus presentation and timing of TMS pulses in one trial in the experimental task. After a fixation cross was presented for 500–1,000 ms, one of four stimuli was presented, toward which participants could respond by pressing the left or right button on a response pad within 1,000 ms. The stimulus was only shown for the first 500 ms. After response, the respective choice was highlighted on screen for 200 ms, followed by 500 ms of blank screen. Positive (“+20 ct”) or negative feedback (“−10 ct”) was then presented on screen for 1,000 ms. Participants needed to learn by trial and error whether one of the choices was related to a higher chance of positive/negative feedback depending on stimulus. Feedback for two of the stimuli had an 80% contingency, while for the other two, it had a 50% contingency. TMS stimulation was applied either 100 ms poststimulus presentation or 100 ms prefeedback stimulation. A total of 336 trials were used in the task.

block was added to increase the number of postlearning trials. This was the case for three participants in one condition and for one participant in both conditions.

#### TMS application and EMG recording

The complete experimental setup is depicted in Figure 4A. Stimulation was applied at 120% of motor threshold (MT) as measured in the first session. MT was measured again on the second session. While there was a trend for a lower motor threshold on the second ( $M = 36.8\%$ ,  $SD = 7.4\%$ ) compared with the first session ( $M = 37.8\%$ ,  $SD = 7.4\%$ ;  $t_{(22)} = 1.72$ ,  $p = 0.100$ ), there was no significant difference in MT between cerebellar and vertex stimulation session ( $t_{(22)} = 0.44$ ,  $p = 0.663$ ).

MT was determined as the lowest intensity that still triggered a motor-evoked potential in at least 5 of 10 stimulations. MEPs were recorded by AgCl surface electrodes (Ambu) from the left M. abductor pollicis brevis in resting condition. The signal was amplified with a Digitimer D360 (Digitimer). The frequency band of the filter was set to 100–5,000 Hz and digitized at a sampling rate of 5 kHz (Signal version 6.02, Cambridge Electronic Design). We monitored for MEPs during the experimental task as to avoid stimulating too close to the brainstem.

TMS was applied via a Magstim Double Cone Coil using a Magstim BiStim<sup>2</sup> unit (Magstim). To enable a fast-paced task flow, we alternated stimulation between two Bistim units. Stimulation was applied either to the left lateral cerebellum (1 cm below and 3 cm to the left of theinion; confer Hardwick et al., 2014) or vertex (at electrode position Cz, Jung et al., 2016), both with inferior voltage flow. The coil was wrapped in plastic wrap to reduce electrode motion artifacts caused by direct contact between TMS and EEG. Participants were given earplugs to reduce auditory artifacts. After the coil was positioned, we fixed it with a custom stand and to the participant's head via a fabric elastic band over the participant's forehead (cerebellar TMS) or chin (vertex TMS). Coil position was constantly monitored and adjusted during the breaks if necessary.

Vertex was chosen as a control site, as it is common choice of control site (Gatti et al., 2023). We did not use sham cerebellar TMS as it provides participants with a very different experience in terms of vibrations, coil clicks, and magnetic field build-up (Duecker and Sack, 2015).

Single-pulse TMS was chosen over rTMS due to its advantage of examining effects of deficits with a high temporal resolution. Instead of examining processing and task performance with a relatively steady deficit across all processing stages, like with rTMS, single-pulse TMS can be applied at different time points within the trial. This offers the advantage of differentiating effects stemming from deficits at different processing stages.

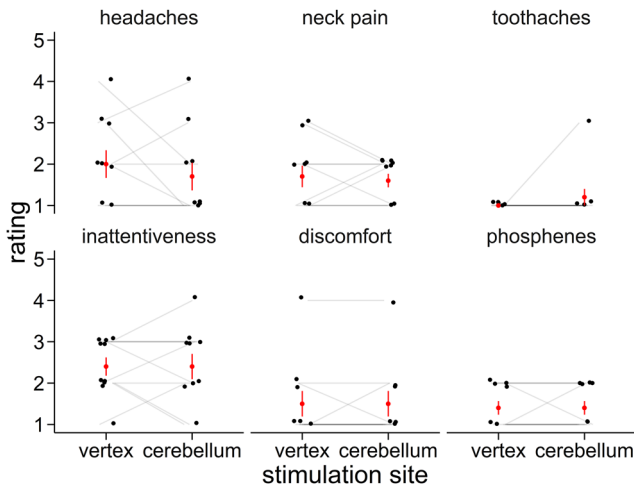
#### Side effects questionnaire

As participants spontaneously reported side effects after the experiment, we introduced a postexperimental side effect questionnaire halfway through the study in which participants were asked to rate symptoms (headaches, neck pain, toothaches, inattentiveness, discomfort, phosphenes, others) associated with TMS on a scale from 1 to 5 (see Fig. 5 for a plot of the side effect ratings). Ten participants completed this questionnaire in both of their sessions and five participants completed it in one session. There were no significant differences between vertex and cerebellar TMS (all  $p \geq 0.343$ ) in terms of reported side effects.

#### EEG recording and preprocessing

All EEG equipment used was explicitly suitable for concurrent TMS. EEG was recorded from 30 passive Ag/AgCl Multitrode electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CPz, CP2, CP6, P7, P3, Pz, P4, P8, O1, Oz, O2, Iz) positioned on a BrainCap (Brain Products) according to the 10–20 system. FCz was used as an on-line reference, and AFz was used as ground electrode. Impedances were kept below 5 k $\Omega$ . Data were amplified with a BrainAmp MR amplifier and recorded at 1,000 Hz using BrainVision Recorder 1.21 (Brain Products).





**Figure 5.** Side effects reported in the postexperimental questionnaire in Experiment 2 (TMS study). Means and standard errors are shown in red, individual ratings are shown in black.

Preprocessing was conducted using the ARTIST algorithm by Wu et al. (2018; see Bertazzoli et al., 2021 for a comparison of TMS-EEG preprocessing methods) and Brain Vision Analyzer. Nevertheless, a remainder of the TMS pulse artifact was still seen in the ERP.

Data were initially checked for missing TMS pulses (TMS marker was sent but pulse was not) both by visual inspection and via an explorative artifact detection: first, trials were segmented around the TMS marker (starting 100 ms before and ending 100 ms after it). Since TMS pulses cause large spikes in the raw data, an automatic artifact detection was employed on the ERP data to identify whether a pulse was sent. Segments with an amplitude  $>400$  or less than  $-400$   $\mu\text{V}$  at electrode Fz were considered to contain a TMS artifact. For all except one participant, all segments contained this artifact, meaning that for each TMS marker in the ERP, a TMS pulse was triggered. Segments in which the above criterion was not met were visually confirmed to contain no TMS pulse artifact and were subsequently excluded from analysis. Marker timings for delayed markers (due to port conflicts) were adjusted in the marker files. This was the case for one marker in 10 participants and two markers in one participant. In two participants, one of these markers indicated a TMS pulse, thus also indicating delay of the corresponding TMS trigger. We excluded the corresponding segments as the TMS pulse had thus not been sent at the correct time.

For preprocessing, we used the ARTIST algorithm by Wu et al. (2018), which is based on EEGLAB (v2022.1; Delorme and Makeig, 2004). This algorithm is designed to decrease artifacts in the EEG signal caused by TMS pulses. In a first step, the ARTIST algorithm corrected for direct current drift, removed the TMS pulse artifact by interpolating the EEG signal around the TMS marker (here: 15 ms prior to until 5 ms after the TMS marker), and removed the decay artifact via ICA. Data were then notch-filtered (50 Hz) and bandpass filtered (high-pass filter: 1 Hz; low-pass filter: 30 Hz). Next, data were segmented into epochs beginning 1,500 ms before and ending 2,200 ms after the TMS markers. Following this, segments containing movement artifacts were rejected ( $M = 2.8\%$  of segments,  $SD = 2.6\%$ ) and bad channels were interpolated ( $M = 0.96$  channels,  $SD = 1.15$  channels). In a final step, bad independent components were removed via a second ICA, after which the signal was rereferenced to an average reference. Deviating from the ARTIST algorithm, we restored electrode FCz after this, because FCz was essential for our data analysis. For baseline correction, the time window between 300 and 100 ms preceding the TMS pulse was used to avoid confounding the baseline correction with the TMS pulse deflection. We always excluded electrode Iz before any preprocessing because it was particularly noisy in pilot testing. Data were then saved in the BrainVision exchange format.

The segment size needed to be rather large, because initial segments were created around the TMS pulse, while in a later step the ERPs needed to be time locked to feedback onset. Therefore, some markers existed in more than one segment (when they were overlapping). To correct this, marker files were edited with a custom MATLAB script, deleting all excess markers. In this step, segments with more than one response were also excluded.

Data were then further preprocessed in BrainVision Analyzer 2.2. Due to only 10 participants exceeding the learning criterion of 75% correct responses in both conditions, the planned data analysis using ANOVA was not feasible. As a result, we pursued a single-trial analysis approach in parallel to data analysis for Experiment 1 and thus deviated from the preregistered procedures. We segmented data around feedback onset, starting 200 ms before and ending 500 ms after feedback markers. Next, we performed an additional baseline correction using the time window from 200 to 0 ms before feedback onset (thus including parts of the remaining pulse artifact for those pulses that were applied 100 ms before the feedback). We then exported single-trial ERPs with a generic data export, on average resulting in 329.3 segments ( $SD = 18.8$  segments) per participant. Data were then exported via a generic data export for further processing in MATLAB. We additionally averaged data according to conditions (stimulation site, TMS timing, feedback valence) to extract FRN peak latencies. Only trials with contingent feedback were included. Peak detection was performed in parallel to Experiment 1.

#### Prediction error estimation

Prediction errors were again modeled as described in Experiment 1. Before merging the behavioral, RL-PE, and EEG data, we excluded all trials in the behavioral and RL-PE data that were not included in the preprocessed EEG data. These were either trials that did not enter the segmentation in ARTIST because the TMS marker/trigger had not been sent (e.g., when participants did not respond in time and thus no feedback-locked TMS trigger was sent) or trials/segments that ARTIST excluded during artifact rejection. Behavioral, RL-PE, and EEG data were then merged.

#### Experimental design and statistical analysis

The study was preregistered to OSF (<https://osf.io/a24rg>). We had aimed for a sample size of 20–25 participants (see preregistration for more details). The targeted sample size was thus matched ( $n = 24$ ). Raw data and code used for preprocessing and analysis are available from <https://osf.io/9n7yp>.

Data were again analyzed in R (version 4.2.3; R Core Team, 2023) using RStudio (version 2023.3.0.386; Posit Team, 2023). Concerning choice accuracy, the preregistered ANOVA analysis as well as an additional linear mixed effects (LME) analysis were performed (see below). Since only 10 participants exceeded the learning criterion of  $>75\%$  correct responses in at least one block for both stimulation sites, the preregistered ANOVA with learning (pre-/postlearning) as a factor was not possible for the FRN analysis. We again decided to pursue a single-trial-based analysis approach using LME models including the unsigned RL-PE instead.

Analyses were conducted to match procedures in Experiment 1, only deviating within the Cook's distance criterion for the choice switching LME analysis where the original criterion was not applicable, so that we instead used the criterion of  $4/n$  (Nieuwenhuis et al., 2012).

**Choice accuracy.** We conducted a repeated-measures ANOVA with the within-subjects factors stimulation site (cerebellum, vertex), TMS timing (poststimulus, prefeedback), and block (1–6), as preregistered. Significant effects were followed up with Bonferroni-corrected  $t$  tests using the function *emmeans\_test*. No participant exceeded the outlier criterion of  $M \pm 2.5$  SD per stimulation site/study session.

**Choice switching.** We analyzed whether choice switching was influenced by the categorical fixed effects feedback valence ( $-0.5$ : negative,  $0.5$ : positive), response type ( $-0.5$ : false,  $0.5$ : correct), stimulation site ( $-0.5$ : vertex,  $0.5$ : cerebellum), TMS timing ( $-0.5$ : poststimulus,  $0.5$ : prefeedback), and the continuous effect block which was scaled via the built-in *scale* function. We also included all interactions of these factors as fixed effects. Three participants had to be excluded

because they exceeded the Cook's distance criterion. The model equation was as follows:

$$\text{choice switching} \sim 1 + \text{feedback valence} * \text{response type} * \text{stimulation site} \\ * \text{TMS timing} * \text{block} + (1 + \text{feedback valence}: \\ \text{response type:stimulation site:TMS timing:block}|\text{subject}).$$

**FRN.** For FRN amplitudes, we again employed LME models with the fixed effects feedback valence (negative:  $-0.5$ , positive:  $0.5$ ), stimulation site ( $-0.5$ : vertex,  $0.5$ : cerebellum), TMS timing ( $-0.5$ : poststimulus,  $0.5$ : prefeedback), and the continuous fixed effect unsigned RL-PE which was the absolute of the signed RL-PE minus  $0.5$  (thus with minimal values of  $-0.5$  and maximal values of  $0.5$ ). We also included all interactions of these factors as fixed effects. While we were initially also able to keep random slopes up to third-level interactions, solving convergence issues via changing the optimizer to *bobyqa*, due to singular fit after the subsequent exclusion of one Cook's distance outlier, we had to revert to a random effects structure with only the fourth-level interactions and random intercept. Two outliers identified by Cook's distance were excluded. The model equation was as follows:

$$\text{FRN} \sim 1 + \text{unsigned PE} * \text{feedback valence} * \text{TMS condition} * \text{TMS timing} \\ * \text{learnability} + (1 + \text{unsigned PE:feedback valence:TMS condition}: \\ \text{TMS timing:learnability}|\text{subject}).$$

## Results

### Experiment 1: RL-PE processing in cerebellar stroke patients

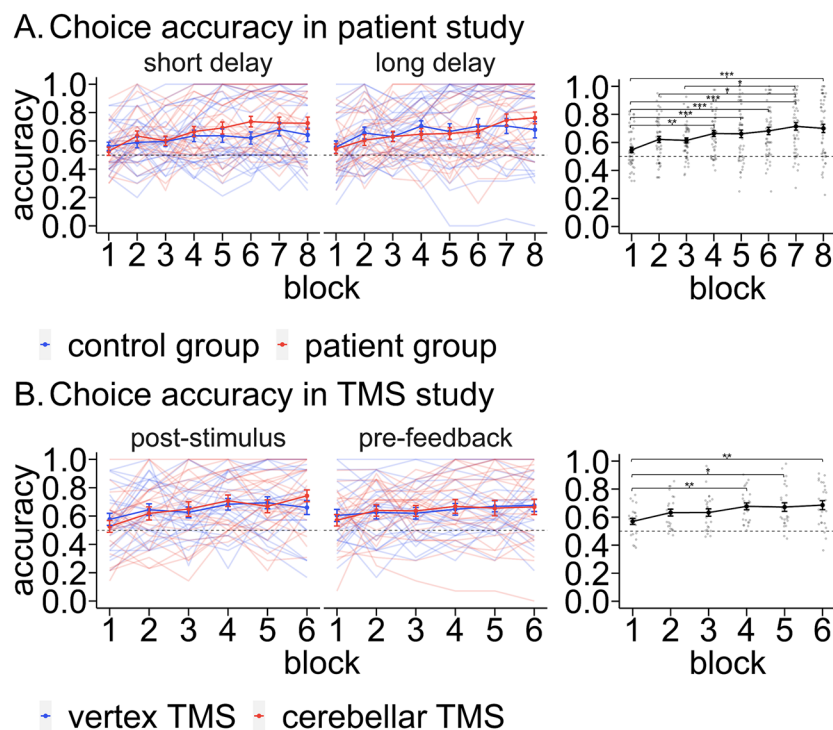
In Experiment 1, we studied patients with chronic cerebellar stroke ( $n = 26$ ) and a matched healthy control group ( $n = 26$ ) to

investigate reinforcement learning success and RL-PE processing, quantified by ACC-driven FRN amplitude. Participants performed a probabilistic feedback learning task (Fig. 2A,B) in which they had to optimize their behavior via trial and error to obtain a monetary reward (“+20 ct” per trial) and avoid a monetary punishment (“−10 ct”). Responses were made by pressing one of two buttons on a response pad. Two out of four stimuli were associated with a 90% reward contingency (i.e., pressing the “correct” button resulted in a reward in 90% of the time and a punishment 10% of the time; vice versa for the “incorrect” button), while the other two stimuli were associated with random feedback. As learning was only possible for the 90% contingency stimuli, the analysis was restricted to these (see Materials and Methods for a more detailed description of the task). Two different feedback delays were used (short, 500 ms; long, 6,500 ms), as previous work has shown differences in FRN depending on feedback timing (Peterburs et al., 2016).

During task performance, EEG was recorded to analyze the FRN (Fig. 2A). To see whether potential deficits were associated with specific lesion locations, lesion symptom mapping was conducted based on T1-weighted MR images in patients.

#### Choice accuracy

Mean choice accuracy by group (patients, controls), feedback delay (short delay, long delay), and block (1–8) is shown in Figure 6A. The effect of these factors on choice accuracy was analyzed within an ANOVA. We expected no differences in accuracy between groups and only a general learning effect. This expectation was based on a previous study which did not find deficient learning



**Figure 6.** Accuracy for Experiments 1 (patient study) and 2 (TMS study). **A**, Left, Choice accuracy in the probabilistic feedback task in Experiment 1 (patient study) according to group (patients, controls), feedback delay (short delay, long delay), and block (1–8). Red lines denote patients and blue lines controls. Opaque lines denote group means. Error bars indicate standard errors. Translucent lines denote individual mean accuracy. Right, Choice accuracy according to block. Opaque lines denote means across groups and feedback delays. Error bars indicate standard errors. Translucent dots denote individual mean accuracy. **B**, Left, Choice accuracy in the probabilistic feedback task in Experiment 2 (TMS study) according to stimulation site (cerebellum, vertex), TMS timing (poststimulus, prefeedback), and block (1–6). Red lines denote vertex TMS and blue lines cerebellar TMS. Opaque lines denote means per stimulation site. Error bars indicate standard errors. Translucent lines denote individual mean accuracy per stimulation site. Right, Choice accuracy according to block. Opaque lines denote means across TMS timings and stimulation sites. Error bars indicate standard errors. Translucent dots denote individual mean accuracy. \* $p < 0.05$ . \*\* $p < 0.01$ . \*\*\* $p < 0.001$ .

in patients with cerebellar stroke (Rustemeier et al., 2016), which is likely due to compensatory mechanisms in this patient group (Peterburs et al., 2012). Overall, we found a significant main effect of block, i.e., a learning effect, but none of the effects involving the factor group (patients vs controls) reached significance.

Statistical analysis showed a main effect of block ( $F_{(3.15,154.26)} = 15.65$ ,  $p < 0.001$ ), indicating that subjects had learned to optimize their behavior over the course of the task. Post hoc  $t$  tests revealed that choice accuracy was significantly higher in block 4 ( $t_{(823)} = 3.97$ ,  $p = 0.002$ ), block 5 ( $t_{(823)} = 3.91$ ,  $p < 0.001$ ), block 6 ( $t_{(823)} = 4.61$ ,  $p < 0.001$ ), block 7 ( $t_{(823)} = 5.72$ ,  $p < 0.001$ ), and block 8 ( $t_{(823)} = 5.26$ ,  $p < 0.001$ ) compared with that in block 1. Additionally, choice accuracy was higher in block 7 compared with that in block 2 ( $t_{(823)} = 3.18$ ,  $p = 0.042$ ) and block 3 ( $t_{(823)} = 3.42$ ,  $p = 0.018$ ). All other pairwise comparisons were nonsignificant (all  $p \geq 0.087$ ). No other effects reached significance (all  $p \geq 0.179$ ; see Table 4 for the complete inferential statistics).

### Choice switching

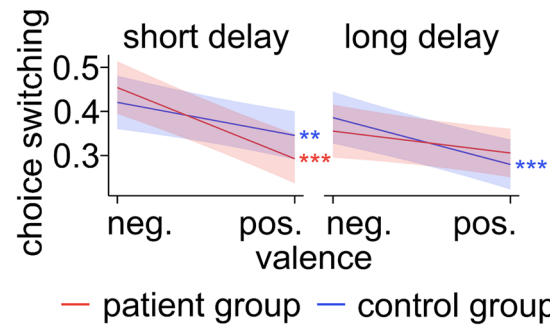
The effects of response type, feedback valence, block, group, and feedback delay on choice switching were analyzed using LME analysis. We found the expected effects of increased choice switching after negative feedback, false responses, short feedback delays, as well as early in the experiment. Importantly, while in controls, choice switching was increased for negative compared with positive feedback for both short and long feedback delays, in patients, this effect was present only for short but not for long feedback delay (Fig. 7A).

Statistical analysis showed that choice switching was increased after incorrect compared with correct responses ( $\beta = -0.23$ ,  $SE = 0.03$ ,  $t_{(15,948.01)} = 8.65$ ,  $p < 0.001$ ). The effect of response type was further modulated by block ( $\beta = -0.10$ ,  $SE = 0.03$ ,  $t_{(15,919.59)} = 3.78$ ,  $p < 0.001$ ), such that the differentiation between correct and false responses was stronger late in the task ( $\beta = -0.33$ ,  $SE = 0.04$ ,  $t = 7.79$ ,  $p < 0.001$ ) but already present early in the task ( $\beta = -0.14$ ,  $SE = 0.04$ ,  $t = 3.77$ ,  $p < 0.001$ ).

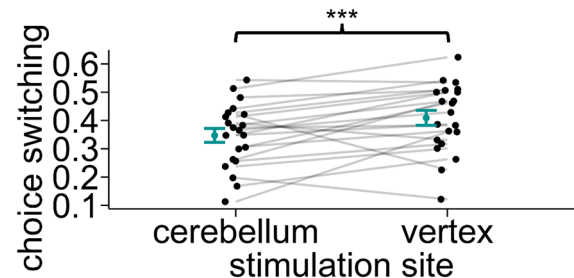
Choice switching was also increased after negative compared with positive feedback ( $\beta = -0.20$ ,  $SE = 0.03$ ,  $t_{(15,920.60)} = 7.60$ ,  $p < 0.001$ ). This effect was further modulated by feedback delay and group ( $\beta = 0.32$ ,  $SE = 0.10$ ,  $t_{(15,826.98)} = 3.10$ ,  $p = 0.002$ ; Fig. 7A). For controls, negative compared with positive feedback resulted in increased choice switching for both short ( $\beta = -0.16$ ,  $SE = 0.05$ ,  $t = 3.27$ ,  $p = 0.004$ ) and long feedback delay ( $\beta = -0.23$ ,  $SE = 0.05$ ,  $t = 4.58$ ,  $p < 0.001$ ). For patients, negative compared with positive feedback resulted in increased choice switching for short ( $\beta = -0.34$ ,  $SE = 0.05$ ,  $t = 6.97$ ,  $p < 0.001$ ) but not for long feedback delay ( $\beta = -0.11$ ,  $SE = 0.05$ ,  $t = 2.13$ ,  $p = 0.134$ ).

Choice switching was also reduced for the long compared with the short feedback delay ( $\beta = -0.08$ ,  $SE = 0.03$ ,  $t_{(15,880.70)} = 3.21$ ,

### A. Choice switching effect in patient study



### B. Choice switching effect in TMS study



**Figure 7.** Choice switching results for Experiments 1 (patient study) and 2 (TMS study). **A**, Slope estimates for choice switching predicted by feedback valence and modulated by feedback delay and group in Experiment 1 (patient study). pos., positive feedback valence; neg., negative feedback valence. Red lines denote patients and blue lines controls. Colored bands indicate 95% confidence intervals. \* $p < 0.05$ . \*\* $p < 0.01$ . \*\*\* $p < 0.001$ . **B**, Mean choice switching according to stimulation site in Experiment 2 (TMS study). Means per stimulation site are displayed in cyan while individual means per stimulation site are displayed in black. Error bars indicate standard errors.

$p = 0.001$ ) and across blocks ( $\beta = -0.08$ ,  $SE = 0.01$ ,  $t_{(15,896.82)} = 5.82$ ,  $p < 0.001$ ). Complete inferential statistics can be found in Table 5.

### FRN

The effects of (signed) RL-PE (reflected by the factors unsigned RL-PE and feedback valence), group, and feedback delay on FRN amplitude were analyzed using LME analysis. We expected the FRN to be increased for high compared with low unsigned RL-PEs for negative feedback and decreased for high compared with low unsigned RL-PEs for positive feedback and expected this effect to be reduced in the patient group. Grand averages of the feedback-locked ERP at FCz show that the FRN amplitude was increased (i.e., more negative) for high compared with low RL-PEs for negative feedback for controls but not for patients (Fig. 8A). This effect could also be confirmed in statistical analysis. For grand average feedback-locked ERPs for all conditions, see Figure 9A.

Statistical analysis showed that the FRN was enhanced for negative compared with positive feedback ( $\beta = 0.48$ ,  $SE = 0.11$ ,  $t_{(1,535.59)} = 4.15$ ,  $p < 0.001$ ) and for high compared with low RL-PEs ( $\beta = -0.39$ ,  $SE = 0.18$ ,  $t_{(769.27)} = 2.16$ ,  $p = 0.031$ ). Importantly, within an interaction between RL-PE, feedback valence, and group ( $\beta = 3.28$ ,  $SE = 1.34$ ,  $t_{(37.08)} = 2.44$ ,  $p = 0.020$ ), the FRN reflected the RL-PE for controls only in negative ( $\beta = -2.10$ ,  $SE = 0.44$ ,  $t = 4.78$ ,  $p < 0.001$ ), but not positive feedback contexts ( $\beta = 0.73$ ,  $SE = 0.39$ ,  $t = 1.88$ ,  $p = 0.242$ ; see Fig. 8B for simple slope plots). For patients, the RL-PE was not reflected for either feedback valence, both  $p > 0.999$ .

**Table 4.** Inferential statistics for the ANOVA investigating the influence of group, feedback delay, and block on accuracy in Experiment 1 (patient study)

Effect	df <sub>n</sub>	df <sub>d</sub>	F	p
Group	1.00	49.00	0.28	0.600
Feedback delay	1.00	49.00	0.50	0.484
Block	3.15	154.26	15.65	<0.001
Group × feedback delay	1.00	49.00	0.96	0.331
Group × block	3.15	154.26	1.64	0.179
Feedback delay × block	4.18	204.87	0.24	0.923
Group × feedback delay × block	4.18	204.87	1.24	0.295

$n = 52$ .



**Table 5. Inferential statistics for the LME analysis examining the effect of feedback valence, response type, feedback delay, group, and block onto choice switching in Experiment 1 (patient study)**

Fixed effects					
	Est/ $\beta$	SE	df	<i>t</i>	<i>p</i>
(Intercept)	0.08	0.04	58.56	1.96	0.055
Feedback valence	−0.20	0.03	15,920.60	−7.60	<0.001
Response type	−0.23	0.03	15,948.01	−8.65	<0.001
Feedback delay	−0.08	0.03	15,880.70	−3.21	0.001
Group	0.00	0.08	58.56	−0.02	0.988
Block	−0.08	0.01	15,896.82	−5.82	<0.001
Feedback valence × response type	−0.07	0.05	15,895.70	−1.26	0.208
Feedback valence × feedback delay	0.09	0.05	15,826.98	1.77	0.076
Response type × feedback delay	−0.10	0.05	15,926.76	−1.92	0.054
Feedback valence × group	−0.05	0.05	15,920.60	−0.87	0.387
Response type × group	−0.08	0.05	15,948.01	−1.45	0.148
Feedback delay × group	0.03	0.05	15,880.70	0.61	0.544
Feedback valence × block	0.02	0.03	15,528.96	0.61	0.542
Response type × block	−0.10	0.03	15,919.59	−3.78	<0.001
Feedback delay × block	0.05	0.03	15,885.45	1.90	0.057
Group × block	0.04	0.03	15,896.82	1.67	0.095
Feedback valence × response type × feedback delay	−0.05	0.10	15,910.73	−0.46	0.648
Feedback valence × response type × group	0.09	0.10	15,895.70	0.85	0.396
Feedback valence × feedback delay × group	0.32	0.10	15,826.98	3.10	0.002
Response type × feedback delay × group	−0.11	0.10	15,926.76	−1.01	0.312
Feedback valence × response type × block	0.05	0.05	15,914.81	1.03	0.305
Feedback valence × feedback delay × block	−0.03	0.05	12,591.80	−0.62	0.535
Response type × feedback delay × block	−0.02	0.05	15,922.19	−0.43	0.670
Feedback valence × group × block	−0.03	0.05	15,528.96	−0.61	0.540
Response type × group × block	−0.06	0.05	15,919.59	−1.12	0.262
Feedback delay × group × block	0.05	0.05	15,885.45	1.02	0.309
Feedback valence × response type × feedback delay × group	−0.13	0.21	15,910.73	−0.64	0.522
Feedback valence × response type × feedback delay × block	−0.12	0.11	322.87	−1.09	0.276
Feedback valence × response type × group × block	−0.15	0.10	15,914.81	−1.44	0.149
Feedback valence × feedback delay × group × block	−0.16	0.10	12,591.80	−1.49	0.136
Response type × feedback delay × group × block	0.14	0.10	15,922.19	1.31	0.190
Feedback valence × response type × feedback delay × group × block	−0.12	0.22	322.87	−0.54	0.587
Random effects					
	Variance	SD	Corr		
Subject (intercept)	0.07	0.26			
Subject (feedback valence × response type × feedback delay × block)	0.05	0.23	−0.62		
Residual	0.86	0.93			
	Marginal		Conditional		
<i>R</i> <sup>2</sup>	0.05		0.12		

Key: *p* values for fixed effects calculated using Satterthwaite's approximations. Model equation: choice switch  $\sim 1 +$  feedback valence \* response type \* feedback delay \* group \* block + (1 + feedback valence:response type:feedback delay:block | subject).  $n_{\text{subjects}} = 52$ ,  $n_{\text{observations}} = 16,001$ .

Within an interaction between feedback valence and feedback delay ( $\beta = -1.09$ ,  $SE = 0.27$ ,  $t_{(73.82)} = 4.04$ ,  $p < 0.001$ ), the FRN was enhanced for negative compared with positive feedback only for short ( $\beta = 0.89$ ,  $SE = 0.16$ ,  $t = 5.66$ ,  $p < 0.001$ ) but not long feedback delay ( $\beta = -0.06$ ,  $SE = 0.16$ ,  $t = 0.37$ ,  $p > 0.999$ ).

It was also increased within a feedback delay main effect for long over short feedback delay ( $\beta = -0.94$ ,  $SE = 0.13$ ,  $t_{(2,765.36)} = 7.24$ ,  $p < 0.001$ ). This effect was further modulated by group within an interaction ( $\beta = 0.59$ ,  $SE = 0.26$ ,  $t_{(2,765.36)} = 2.27$ ,  $p = 0.024$ ). The FRN amplitude was more strongly increased for long feedback delays for controls ( $\beta = -1.51$ ,  $SE = 0.19$ ,  $t = 7.93$ ,  $p < 0.001$ ) than patients ( $\beta = -0.68$ ,  $SE = 0.16$ ,  $t = 4.15$ ,  $p < 0.001$ ).

Complete inferential statistics can be found in Table 6.

### Lesion symptom mapping

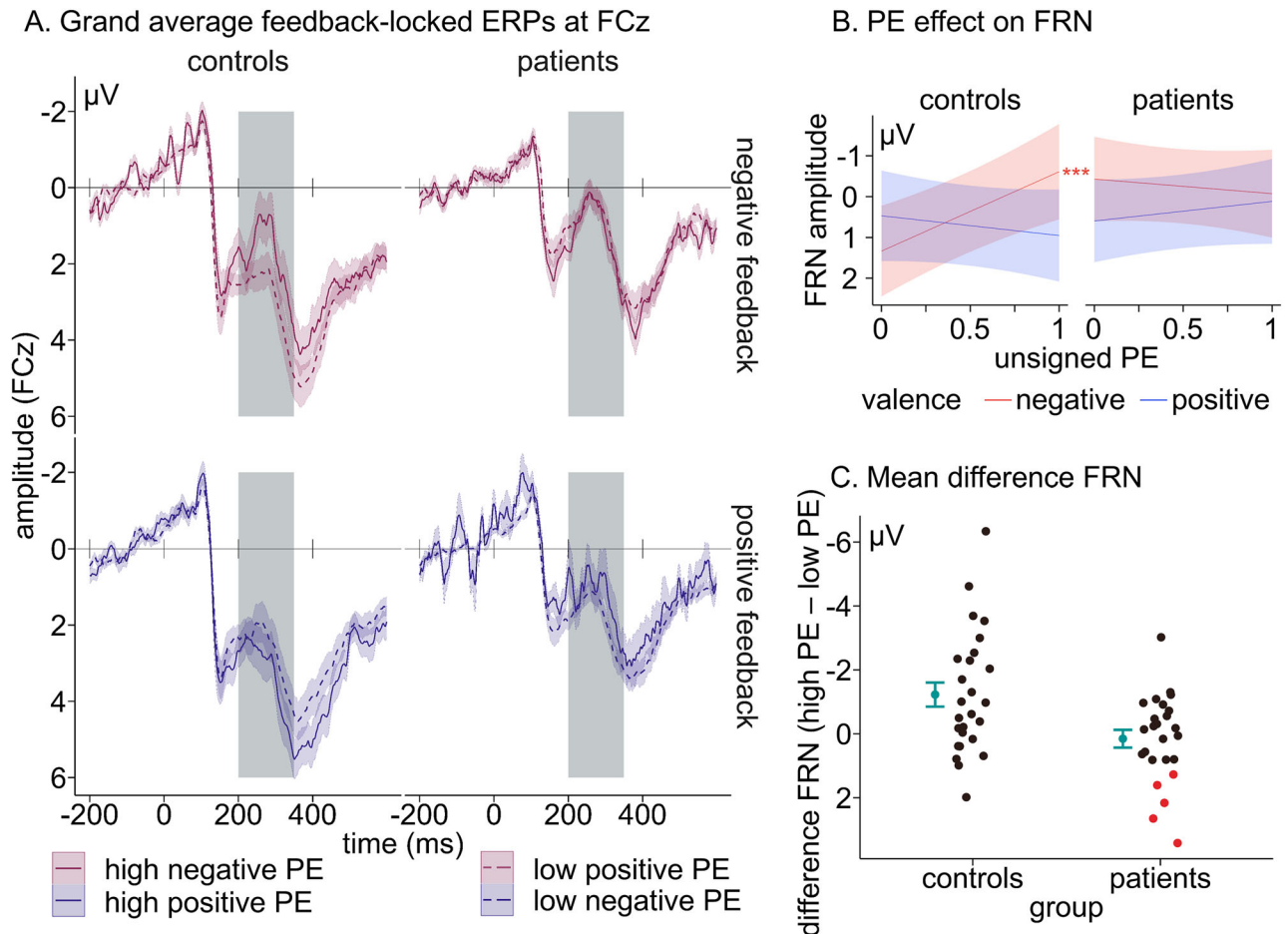
Lesions were mainly located in posterolateral regions, with highest overlap in lobules Crus II, I, VIIb, VIIa, and VIIb (Fig. 1). Six patients had a lesion extending into the dentate nucleus. Images of individual lesions are displayed in Extended Data Figure 1-1. For analysis, all lesions were mirrored to the left side (in case of bilateral lesions, the side with the larger lesion was mirrored to the left side if necessary).

To investigate whether the FRN changes were linked to specific cerebellar lesion locations, voxel-based lesion symptom mapping (vblSM) was performed. While controls mostly showed the expected coding of RL-PEs in the FRN in the anticipated direction (i.e., increased/more negative FRN amplitude for high over low RL-PEs), only few patients showed this pattern, and some patients even showed the opposite (i.e., decreased/more positive FRN for high over low RL-PE; Fig. 8C). We used the difference FRN for negative feedback as a parameter for the vblSM (FRN for high RL-PEs  $\geq 0.5$  – FRN for low RL-PEs  $< 0.5$ ). We expected aberrant processing to be associated with damage to posterolateral regions, especially Crus I and II. Indeed, a more aberrant difference FRN was associated with more four posterior lesion clusters: in Crus II extending toward lobule VIIb (peak  $z = 3.0$ , peak coordinates:  $x = -26$  mm,  $y = -78$  mm,  $z = -51$  mm,  $535 \text{ mm}^3$ ), medial Crus II (peak  $z = 2.6$ , peak coordinates:  $x = -5$  mm,  $y = -79$  mm,  $z = -35$  mm,  $37 \text{ mm}^3$ ), Crus I (peak  $z = 2.5$ , peak coordinates:  $x = -27$  mm,  $y = -86$  mm,  $z = -34$  mm,  $149 \text{ mm}^3$ ), and medial lobule VIIb/vermal VIIa (peak  $z = 2.3$ , peak coordinates:  $x = -6$  mm,  $y = -68$  mm,  $z = -45$  mm,  $550 \text{ mm}^3$ ; Fig. 10).

In an additional step, we examined lesions in the dentate nucleus (see Materials and Methods). There were only six patients with lesions in the dentate nucleus (of which two had been classified as impaired) with only minimal overlap. Meaningful analyses could thus not be performed. Plots of individual dentate nucleus lesions are shown in Figure 3.

### Experiment 2: RL-PE processing in healthy young adults receiving cerebellar TMS

In Experiment 2, we investigated reinforcement learning and RL-PE processing in young healthy adults ( $n = 24$ ) for cerebellar and control (vertex) single-pulse TMS using the same probabilistic feedback learning task as in Experiment 1 (Fig. 4A). Pulses were applied once per trial and either at the response stage (100 ms poststimulus onset) or at the feedback stage (100 ms prefeedback; Fig. 4B). Single-pulse TMS has the advantage of transient effects on behavior and neural processing (Gatti et al.,



**Figure 8.** ERP results for Experiment 1 (patient study). **A**, Grand average feedback-locked ERPs at FCz according to unsigned RL-PE (low, high), feedback valence (positive, negative), and group (patients, controls). Red lines denote high unsigned RL-PE ( $>0.5$ ) and blue lines low unsigned RL-PE ( $\leq 0.5$ ). Colored bands indicate standard errors. **B**, Slope estimates for FRN amplitude predicted by unsigned RL-PE and modulated by feedback valence and group. Red lines denote positive feedback valence and blue lines negative feedback valence. Colored bands indicate 95% confidence intervals.  $*p < 0.05$ .  $**p < 0.01$ .  $***p < 0.001$ . **C**, Mean difference FRN (mean negative high RL-PE – mean negative low RL-PE) separately for groups (controls, patients). Group means are displayed in cyan while individual means are displayed in black. Error bars indicate standard errors. Patients with a difference FRN above 1  $\mu V$  (i.e., decreased/more positive difference FRN) are marked in red who are used as impaired group in lesion symptom mapping.

2023) within-subject, excluding the possibility of long-term compensation that may be present in chronic stroke patients (Peterburs et al., 2012).

#### Choice accuracy

Mean choice accuracy as a function of stimulation site (cerebellum, vertex), TMS timing (poststimulus, prefeedback), and block (1–8) is shown in Figure 6B. The effects of these factors on choice accuracy were analyzed using an ANOVA. We expected a main effect of block, and participants to perform worse when receiving cerebellar compared with vertex TMS, as no long-term compensatory mechanisms should be available due to the instantaneous effect of the TMS. Overall, we found a main effect of block, i.e., a learning effect, while no effects involving the stimulation site factor (cerebellum/vertex) reached significance.

Statistical analysis showed a significant main effect of block ( $F_{(3.18,73.05)} = 6.21$ ,  $p < 0.001$ ) with higher choice accuracy in block 4 ( $t_{(570)} = 3.49$ ,  $p = 0.008$ ), block 5 ( $t_{(570)} = 3.33$ ,  $p = 0.014$ ), and block 6 ( $t_{(570)} = 3.77$ ,  $p = 0.003$ ) compared with block 1. All other effects were nonsignificant (all  $p \geq 0.461$ ; see Table 7 for complete inferential statistics).

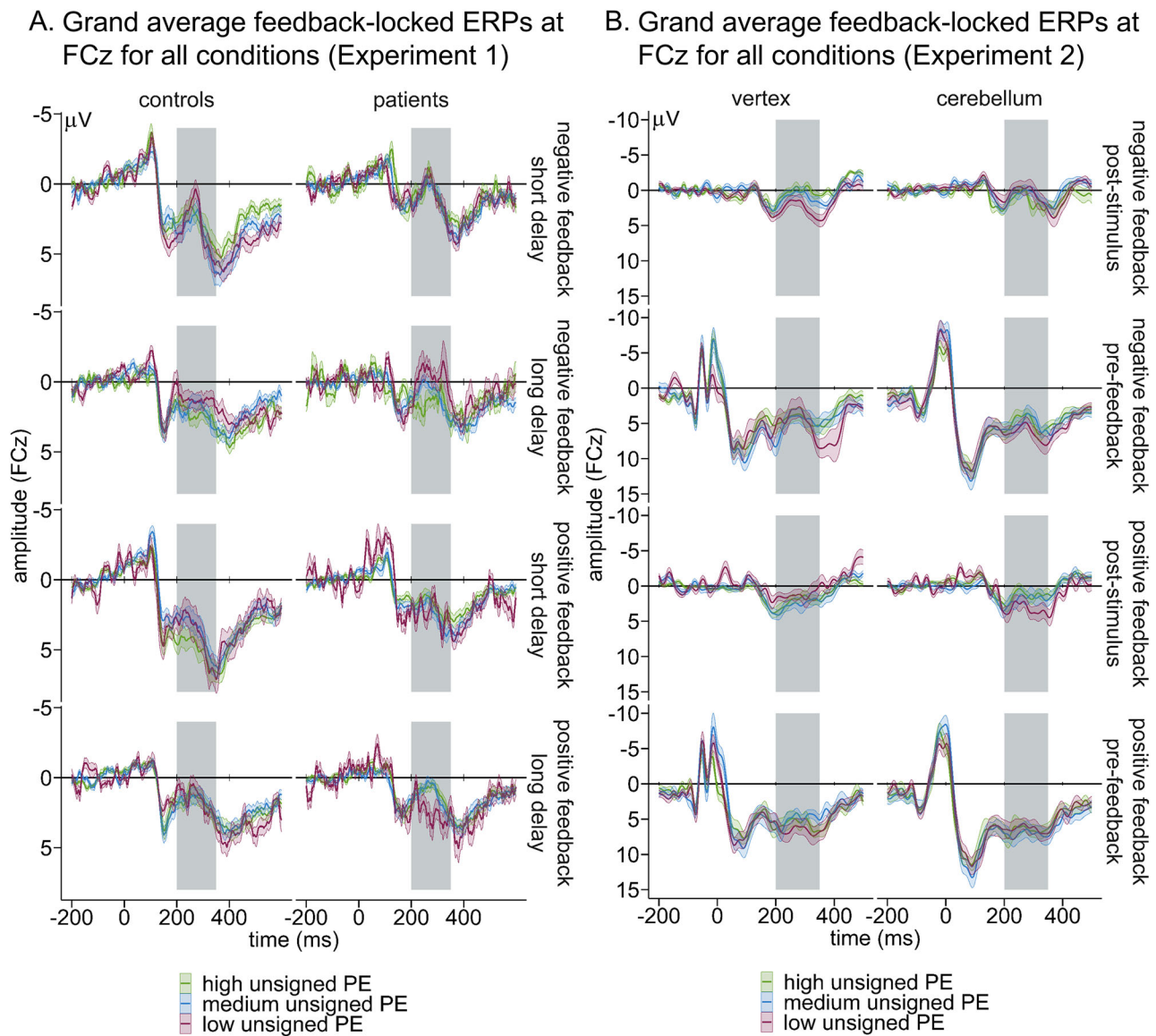
#### Choice switching

The effects of response type, feedback valence, block, stimulation site, and TMS timing on choice switching were analyzed using LME analysis. We found the expected effects of increased choice switching after negative feedback and false responses. Importantly, choice switching was generally reduced for cerebellar compared with vertex stimulation (Fig. 7B).

The main effect of stimulation site reached significance ( $\beta = -0.11$ ,  $SE = 0.03$ ,  $t_{(6,475.16)} = 3.67$ ,  $p < 0.001$ ), with decreased choice switching for cerebellar compared with vertex TMS (Fig. 7B).

Further, the main effect of response type was significant ( $\beta = -0.39$ ,  $SE = 0.03$ ,  $t_{(6,502.49)} = 12.64$ ,  $p < 0.001$ ), with more choice switching after incorrect compared with correct choices. This effect was further modulated by block ( $\beta = -0.15$ ,  $SE = 0.03$ ,  $t_{(6,499.06)} = 5.02$ ,  $p < 0.001$ ). Follow-up simple slope analyses showed that while choice switching was significantly increased both early ( $\beta = -0.22$ ,  $SE = 0.04$ ,  $t = 5.22$ ,  $p < 0.001$ ) and late in the task ( $\beta = -0.54$ ,  $SE = 0.05$ ,  $t = 11.44$ ,  $p < 0.001$ ), the effect was stronger late in the task.

Statistical analysis also showed a main effect of feedback valence ( $\beta = -0.23$ ,  $SE = 0.03$ ,  $t_{(6,489.74)} = 7.74$ ,  $p < 0.001$ ), with more choice switching after negative compared with positive feedback.



**Figure 9.** Grand average feedback-locked ERPs for all conditions and three RL-PE levels. **A**, Grand average feedback-locked ERPs for Experiment 1 (patient study) at FCz according to unsigned RL-PE (low, medium, high), feedback valence (positive, negative), feedback delay (short delay, long delay), and group (patients, controls). Red lines denote low, blue lines denote medium, and green lines denote high unsigned RL-PE. Colored bands indicate standard errors. **B**, Grand average feedback-locked ERPs for Experiment 2 (TMS study) at FCz according to unsigned RL-PE (low, medium, high), feedback valence (positive, negative), TMS timing (poststimulus, prefeedback), and stimulation site (vertex, cerebellum). Red lines denote low, blue lines denote medium, and green lines denote high unsigned RL-PE. Colored bands indicate standard errors.

No other effects involving response type or feedback valence with stimulation site or each other emerged (all  $p \geq 0.246$ ). Complete inferential statistics can be found in Table 8.

#### FRN

The effects of (signed) RL-PE (reflected in unsigned RL-PE and feedback valence), stimulation site, and TMS timing on FRN amplitude were analyzed using LME analysis. We expected the FRN to be increased for higher unsigned RL-PEs for negative feedback and decreased for higher unsigned RL-PEs for positive feedback, and we expected this effect to be reduced when stimulating the cerebellum compared with the vertex. Grand averages for the feedback-locked ERP at FCz are shown in Figure 11A. For grand average feedback-locked ERPs for all conditions, see Figure 9B. FRN amplitudes were reduced for high compared with low

RL-PEs for negative feedback only for control stimulation but not for cerebellar stimulation.

Statistical analysis showed that FRN was enhanced for negative compared with positive feedback ( $\beta = 1.08$ , SE = 0.14,  $t_{(7,116.94)} = 7.90$ ,  $p < 0.001$ ). Feedback valence further interacted significantly with the unsigned RL-PE, ( $\beta = -1.11$ , SE = 0.49,  $t_{(6,506.98)} = 2.28$ ,  $p = 0.023$ ). Follow-up simple slope analyses revealed that while the unsigned RL-PE modulated the FRN for negative feedback, with a reduced FRN with increasing RL-PE at trend level ( $\beta = 0.70$ , SE = 0.32,  $t = 2.18$ ,  $p = 0.058$ ), this was not the case for positive feedback ( $\beta = -0.12$ , SE = 0.29,  $t = 0.42$ ,  $p > 0.999$ ). This interaction was further modulated by stimulation site ( $\beta = 2.82$ , SE = 0.93,  $t_{(5,523.19)} = 3.05$ ,  $p = 0.002$ ; Fig. 11B). For vertex TMS, FRN was reduced with increasing unsigned RL-PE for negative feedback ( $\beta = 1.57$ , SE = 0.43,  $t = 3.69$ ,  $p < 0.001$ ) but not positive feedback ( $\beta = -0.56$ ,



**Table 6. Inferential statistics for the LME analysis examining the effect of unsigned RL-PE, feedback valence, feedback delay, and group onto FRN amplitude in Experiment 1 (patient study)**

Fixed effects					
	Est/ $\beta$	SE	df	<i>t</i>	<i>p</i>
(Intercept)	0.29	0.34	46.00	0.87	0.389
Unsigned PE	−0.39	0.18	769.27	−2.16	0.031
Feedback valence	0.48	0.11	1,535.59	4.15	<0.001
Feedback delay	−0.94	0.13	2,765.36	−7.24	<0.001
Group	−0.48	0.67	46.00	−0.72	0.477
Unsigned PE $\times$ feedback valence	0.79	0.67	37.08	1.18	0.246
Unsigned PE $\times$ feedback delay	0.56	0.47	41.99	1.19	0.239
Feedback valence $\times$ feedback delay	−1.09	0.27	73.82	−4.04	<0.001
Unsigned PE $\times$ group	0.67	0.37	769.27	1.83	0.068
Feedback valence $\times$ group	0.26	0.23	1,535.59	1.13	0.258
Feedback delay $\times$ group	0.59	0.26	2,765.36	2.27	0.024
Unsigned PE $\times$ feedback valence $\times$ feedback delay	−2.02	1.73	36.29	−1.17	0.250
Unsigned PE $\times$ feedback valence $\times$ group	−3.28	1.34	37.08	−2.44	0.020
Unsigned PE $\times$ feedback delay $\times$ group	−0.99	0.94	41.99	−1.06	0.295
Feedback valence $\times$ feedback delay $\times$ group	0.09	0.54	73.82	0.16	0.872
Unsigned PE $\times$ feedback valence $\times$ feedback delay $\times$ group	−4.03	3.45	36.29	−1.17	0.251
Random effects					
	Variance	SD	Corr		
Subject (intercept)	5.20	2.28			
Subject (unsigned PE $\times$ feedback valence)	11.10	3.33	−0.07		
Subject (unsigned PE $\times$ feedback delay)	4.16	2.04	0.42	0.65	
Subject (feedback valence $\times$ feedback delay)	0.96	0.98	−0.07	−0.96	−0.70
Subject (unsigned PE $\times$ feedback valence $\times$ feedback delay)	104.57	10.23	−0.08	−0.52	−0.16
Residual	37.25	6.10			
Model fit					
	Marginal		Conditional		
$R^2$	0.01		0.16		

Key: *p* values for fixed effects calculated using Satterthwaite's approximations. Model equation:  $FRN \sim 1 + \text{unsigned PE} * \text{feedback valence} * \text{feedback delay} * \text{group} + (1 + \text{unsigned PE} : \text{feedback valence} : \text{feedback delay} + \text{unsigned PE} : \text{feedback valence} + \text{unsigned PE} : \text{feedback delay} + \text{feedback valence} : \text{feedback delay}) | \text{subject}$ .  $n_{\text{subjects}} = 48$ ,  $n_{\text{observations}} = 15,034$ .

SE = 0.39,  $t = 1.45$ ,  $p = 0.594$ ). When stimulating the cerebellum, the FRN was not significantly modulated by the unsigned RL-PE (both  $p \geq 0.999$ ).

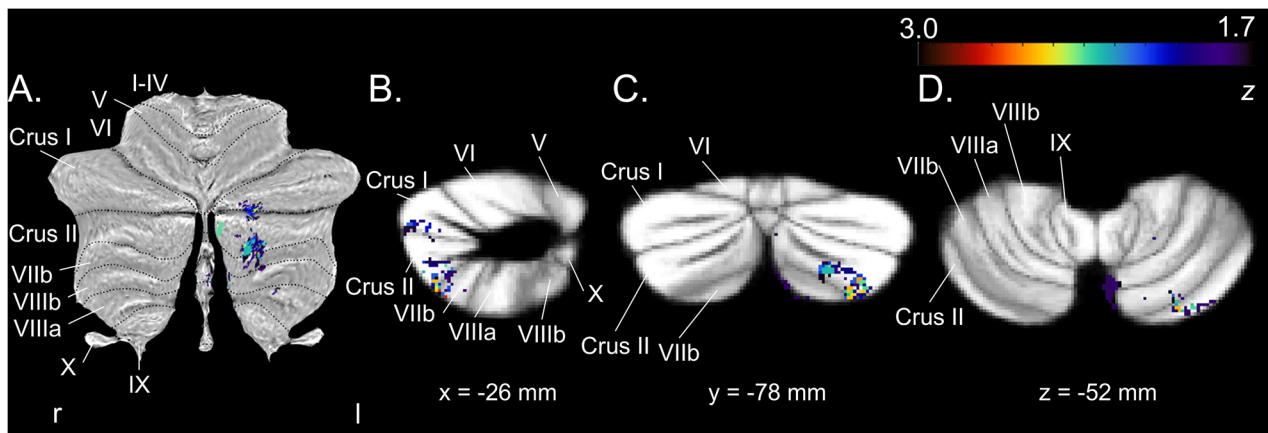
No other effects involving feedback valence or the unsigned RL-PE with stimulation site or each other emerged (all  $p \geq 0.207$ ). Complete inferential statistics can be found in Table 9.

### Control analysis

To explore whether feedback processing was generally disrupted or whether this was more specific to the processing of RL-PEs, we performed a control analysis investigating whether patients with cerebellar stroke (Experiment 1) and healthy controls receiving cerebellar TMS (Experiment 2) showed preserved valence coding in the FRN, as valence effects for short feedback delays are a well-reported finding (Sambrook and Goslin, 2015; Hinneberg and Hegele, 2022). This was investigated within the same LME model reported above, resolving the (nonsignificant) interactions between outcome valence, feedback delay, and group for Experiment 1 and outcome valence, TMS timing, and stimulation site for Experiment 2.

Indeed, result patterns were consistent with intact valence coding in FRN under short feedback delays for patients with cerebellar stroke and healthy participants receiving cerebellar TMS. In Experiment 1, the FRN was indeed more negative for negative relative to positive feedback in both controls ( $\beta = 0.81$ , SE = 0.23,  $t = 3.54$ ,  $p = 0.002$ ) and patients ( $\beta = 1.01$ , SE = 0.21,  $t = 4.79$ ,  $p < 0.001$ ) in the short delay condition but not in the long delay condition (both  $p \geq 0.103$ ; Fig. 12A). In Experiment 2, where only short delays were used, the FRN was consistently more negative for negative over positive valence for both stimulation sites and TMS timings (poststimulus vertex:  $\beta = 1.12$ , SE = 0.27,  $t = 4.13$ ,  $p < 0.001$ ; poststimulus cerebellum:  $\beta = 0.79$ , SE = 0.27,  $t = 2.96$ ,  $p = 0.012$ ; prefeedback vertex:  $\beta = 1.47$ , SE = 0.27,  $t = 5.54$ ,  $p < 0.001$ ; prefeedback cerebellum:  $\beta = 1.09$ , SE = 0.27,  $t = 4.07$ ,  $p < 0.001$ ; Fig. 12B).

To further examine whether we could find evidence against an effect of these interactions, model comparisons were performed to extract an estimation of the Bayes factor. The full LME models were compared against a model without the interaction between outcome valence, feedback delay, and group for Experiment 1 and outcome valence, TMS timing, and stimulation site for Experiment 2. The Bayes factor was then estimated based on the difference Bayesian Information Criterion, following Shen and González (2021). In line with the control analyses, the



**Figure 10.** Lesion symptom mapping in Experiment 1 (patient study). Voxel-based lesion symptom mapping of lesion location comparing groups on (A) a cerebellar flatmap (Diedrichsen and Zotow, 2015) and in 2D (B) sagittal, (C) coronal, and (D) axial views. Color code shown on the top right denotes z-scores (from purple = 1.7 to red = 3.0). r, right; l, left.

**Table 7. Inferential statistics for the ANOVA investigating the influence of stimulation site, TMS timing and block on accuracy in Experiment 2 (TMS study)**

Effect	df <sub>n</sub>	df <sub>d</sub>	F	p
Stimulation site	1	23	0.01	0.938
TMS timing	1	23	0.13	0.717
Block	3.18	73.05	6.21	<0.001
Stimulation site × TMS timing	1	23	0.01	0.929
Stimulation site × block	3.12	71.87	0.88	0.461
TMS timing × block	3.33	76.65	0.62	0.619
Stimulation site × TMS timing × block	5	115	0.65	0.660

n = 24.

estimated Bayes Factor indicated very strong evidence for an absence of the triple interactions (BF = 90.02 for both experiments/comparisons).

**Exploratory analysis of predictability of choice switching by ERP components**

While differences in accuracy or choice switching between groups (Experiment 1) and stimulation sites (Experiment 2) were either not significant or not as severe, differences in RL-PE processing in the FRN were substantial. While the FRN does not seem to have a strong behavioral correlate (Ullsperger, 2024), there is some evidence linking it to behavioral flexibility (Cohen and Ranganath, 2007; Fischer and Ullsperger, 2013; Kirschner et al., 2022; but also see Chase et al., 2011). In an exploratory analysis, we added the FRN amplitude as an additional factor to the LME models with choice switching as the dependent variable. However, the respective LME models did not offer a better fit to the data than the original models (Experiment 1:  $\chi^2_{(32)} = 36.01$ ,  $p = 0.286$ ; Experiment 2:  $\chi^2_{(32)} = 24.82$ ,  $p = 0.813$ ) and neither did any effects including the FRN amplitude and the group factor reach significance within these models (all  $p \geq 0.079$ ). In an additional analysis for Experiment 1, we tried to relate the difference FRN as presented in Figure 8C to choice switching using group-specific correlations. However, this correlation did not reach significance for controls ( $t_{(24)} = 0.29$ ,  $p = 0.773$ ) or patients ( $t_{(23)} = 0.28$ ,  $p = 0.784$ ).

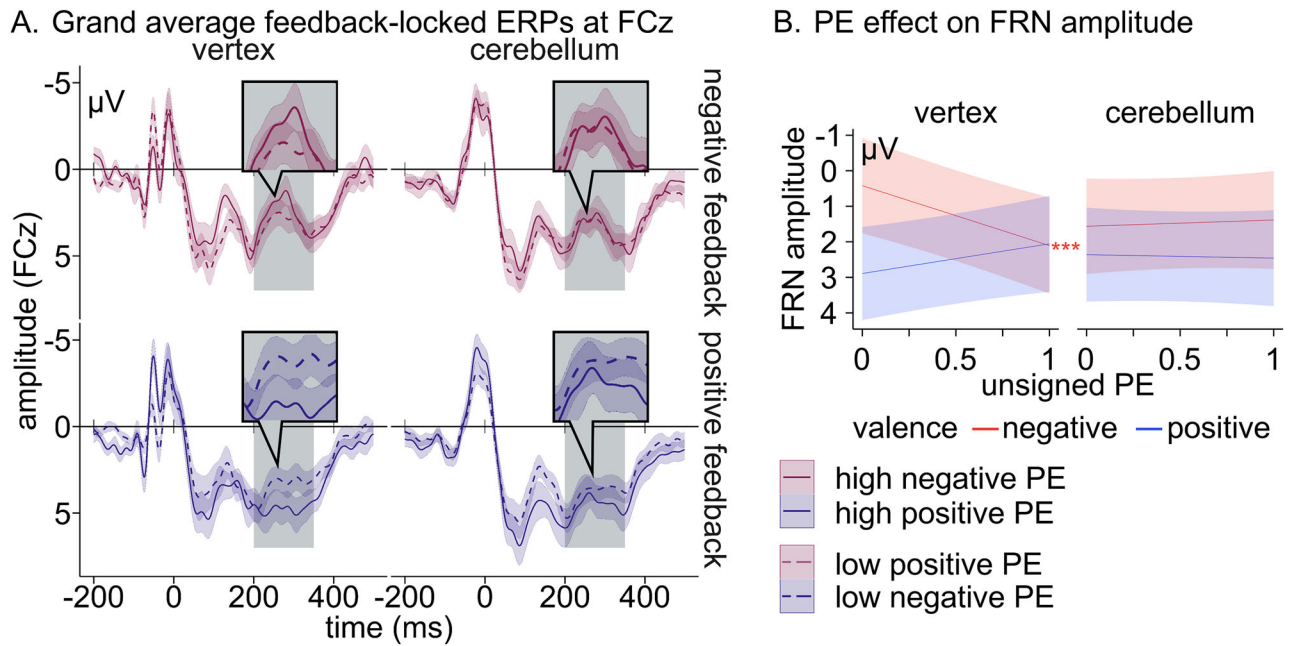
For completeness, we also added the P3a and P3b separately to the choice switching model to see whether these would improve model fit. P3a and P3b in the feedback-related ERP peak at ~300–500 ms postfeedback frontocentrally and parietally, respectively (Hruby and Marsalek, 2003; Polich, 2007), and have been more clearly associated with behavioral flexibility (Ullsperger, 2024). Indeed, including P3a/b in the model significantly improved model fit for both P3a ( $\chi^2_{(32)} = 58.19$ ,  $p = 0.003$ ) and P3b ( $\chi^2_{(32)} = 54.72$ ,  $p = 0.007$ ) in Experiment 1. Full statistics for the model comparisons can be found in Table 10. Several effects including both P3a and group reached significance and one effect including both P3b and group reached significance. The general effect patterns seem to indicate that the P3a is generally predictive of choice switching for healthy controls, but only under specific circumstances for patients.

For the P3a LME model in Experiment 1, the interaction between P3a and group reached significance ( $\beta = 0.01$ , SE = 0.00,  $t_{(15,752.83)} = 2.04$ ,  $p = 0.042$ ). P3a was predictive of choice switching for controls ( $\beta = -0.01$ , SE = 0.00,  $t = -2.45$ ,  $p = 0.028$ ), such that an increased P3a corresponded to decreased choice switching. This effect did not reach significance in patients ( $\beta = 0.00$ , SE = 0.00,  $t = 0.90$ ,  $p = 0.734$ ). This interaction was further modulated by block ( $\beta = 0.01$ , SE = 0.00,  $t_{(15,710.76)} = 2.05$ ,

**Table 8. Inferential statistics for the LME analysis examining the effect of feedback valence, response type, stimulation site, TMS timing, and block onto choice switching in Experiment 2 (TMS study)**

Fixed effects					
	Est/β	SE	df	t	p
(Intercept)	0.07	0.04	22.60	1.91	0.069
Feedback valence	−0.23	0.03	6,489.74	−7.74	<0.001
Response type	−0.39	0.03	6,502.49	−12.64	<0.001
Stimulation site	−0.11	0.03	6,475.16	−3.67	<0.001
TMS timing	0.00	0.03	6,477.69	0.06	0.954
Block	0.00	0.02	6,481.40	0.21	0.831
Feedback valence × response type	0.07	0.06	6,479.45	1.15	0.252
Feedback valence × stimulation site	−0.01	0.06	6,487.20	−0.19	0.852
Response type × stimulation site	−0.06	0.06	6,488.89	−1.07	0.286
Feedback valence × TMS timing	0.01	0.06	6,461.84	0.14	0.889
Response type × TMS timing	0.04	0.06	6,502.24	0.64	0.524
Stimulation site × TMS timing	−0.03	0.06	6,479.64	−0.56	0.577
Feedback valence × block	0.01	0.03	6,327.19	0.36	0.720
Response type × block	−0.15	0.03	6,499.06	−5.02	<0.001
Stimulation site × block	−0.03	0.03	6,482.72	−0.87	0.387
TMS timing × block	0.01	0.03	6,483.30	0.31	0.758
Feedback valence × response type × stimulation site	0.13	0.12	6,481.54	1.04	0.298
Feedback valence × response type × TMS timing	0.04	0.12	6,489.48	0.35	0.727
Feedback valence × stimulation site × TMS timing	−0.12	0.12	6,436.24	−0.99	0.325
Response type × stimulation site × TMS timing	0.02	0.12	6,502.30	0.19	0.848
Feedback valence × response type × block	−0.07	0.06	6,484.86	−1.11	0.268
Feedback valence × stimulation site × block	0.00	0.06	6,107.47	−0.04	0.968
Response type × stimulation site × block	−0.02	0.06	6,490.12	−0.36	0.721
Feedback valence × TMS timing × block	0.13	0.06	6,384.46	2.08	0.037
Response type × TMS timing × block	−0.03	0.06	6,485.44	−0.48	0.629
Stimulation site × TMS timing × block	−0.07	0.06	6,479.98	−1.16	0.246
Feedback valence × response type × stimulation site × TMS timing	−0.03	0.24	6,489.04	−0.12	0.906
Feedback valence × response type × stimulation site × block	−0.04	0.12	6,485.43	−0.37	0.709
Feedback valence × response type × TMS timing × block	−0.27	0.12	6,489.30	−2.22	0.027
Feedback valence × stimulation site × TMS timing × block	0.03	0.12	4,922.33	0.21	0.833
Response type × stimulation site × TMS timing × block	0.13	0.12	6,489.78	1.10	0.271
Feedback valence × response type × stimulation site × TMS timing × block	0.01	0.25	48.64	0.04	0.971
Random effects					
	Variance	SD	Corr		
Subject (intercept)	0.03	0.16			
Subject (feedback valence × response type × stimulation site × TMS timing × block)	0.10	0.32	0.53		
Residual	0.88	0.94			
Model fit					
	Marginal		Conditional		
R <sup>2</sup>	0.08		0.11		

Key: p values for fixed effects calculated using Satterthwaite's approximations. Model equation: choice switching ~ 1 + feedback valence \* response type \* stimulation site \* TMS timing \* block + (1 + feedback valence: response type: stimulation site: TMS timing: block | subject). n<sub>subjects</sub> = 21, n<sub>observations</sub> = 6,541.



**Figure 11.** ERP results for Experiment 2 (TMS study). **A**, Grand average feedback-locked ERPs at FCz according to unsigned RL-PE (low, high), feedback valence (positive, negative), and stimulation site (vertex, cerebellum). Red lines denote high unsigned RL-PE (>0.5) and blue lines low unsigned RL-PE (≤0.5). Colored bands indicate standard errors. **B**, Slope estimates for FRN amplitude predicted by unsigned RL-PE and modulated by feedback valence and stimulation site. Red lines denote positive feedback valence and blue lines negative feedback valence. Colored bands indicate 95% confidence intervals. \* $p < 0.05$ . \*\* $p < 0.01$ . \*\*\* $p < 0.001$ .

$p = 0.041$ ). Choice switching was increased with increasing P3a for patients late in the task on trend level ( $\beta = 0.01$ ,  $SE = 0.00$ ,  $t = 2.38$ ,  $p = 0.068$ ), but not early in the task, and not at all for healthy controls (all  $p \geq 0.191$ ). The interaction was further modulated by outcome valence in a four-way interaction ( $\beta = 0.02$ ,  $SE = 0.01$ ,  $t_{(15,706.65)} = 2.31$ ,  $p = 0.021$ ). Choice switching was increased on trend level with increasing P3a for patients late in the task only for positive feedback ( $\beta = 0.01$ ,  $SE = 0.01$ ,  $t = 2.55$ ,  $p = 0.087$ ). No other simple slopes reached significance when resolving the interaction (all  $p \geq 0.168$ ).

A three-way interaction including P3a, group, and feedback delay ( $\beta = -0.02$ ,  $SE = 0.01$ ,  $t_{(15,729.44)} = 2.23$ ,  $p = 0.026$ ) showed that P3a was predictive of choice switching only for controls in the short feedback delay condition, however, only at trend level ( $\beta = -0.01$ ,  $SE = 0.00$ ,  $t = 2.39$ ,  $p = 0.067$ ). Larger P3a amplitudes led to reduced choice switching. The effect did not reach significance for controls in the long feedback delay condition or for patients at all (all  $p \geq 0.125$ ). However, two higher-level interactions indicated that a P3a-choice switching relation did exist for patients under specific circumstances: The interaction between P3a, group, and feedback delay was further modulated by response type in a four-way interaction ( $\beta = -0.04$ ,  $SE = 0.02$ ,  $t_{(15,712.02)} = 2.37$ ,  $p = 0.018$ ). This four-way interaction revealed that P3a was significantly predictive of choice switching only for patients for short feedback delay when the choice was correct ( $\beta = 0.02$ ,  $SE = 0.00$ ,  $t = 3.48$ ,  $p = 0.004$ ). In no other condition did the effect reach significance (all  $p \geq 0.301$ ). Finally, the four-way interaction was further modulated by outcome valence in a five-way interaction ( $\beta = 0.08$ ,  $SE = 0.04$ ,  $t_{(15,705.28)} = 2.16$ ,  $p = 0.031$ ). In this five-way interaction, P3a was predictive of choice switching on trend level only for patients for the short delay condition, when feedback was positive and the reaction was correct, with larger P3a amplitudes predicting more choice switching ( $\beta = 0.01$ ,  $SE = 0.00$ ,  $t = 2.77$ ,  $p = 0.090$ ). No other slope reached significance (all  $p \geq 0.254$ ). These effect patterns might

be due to a smaller, more general effect of P3a on choice switching in healthy controls, while for patients, the direction of the effect was more dependent on the experimental condition. The full statistical pattern is displayed in Table 11.

For P3b, a four-way interaction between P3b, outcome valence, group, and block emerged ( $\beta = 0.02$ ,  $SE = 0.01$ ,  $t_{(14,135.18)} = 1.97$ ,  $p = 0.049$ ). P3b amplitude was predictive of choice switching only for patients when receiving positive feedback and only late in the task ( $\beta = 0.02$ ,  $SE = 0.01$ ,  $t = 3.97$ ,  $p < 0.001$ ; all other  $p \geq 0.894$ ).

Notably, we could not replicate these findings for Experiment 2. The stroke patient sample might have been more suitable to explore such deficits, as the single-pulse TMS we used in Experiment 2 creates a virtual lesion merely via inducing noise. A brain-behavior connection might be clearer when using rTMS protocols. Including the P3a/b in the model did not result in significant improvements in model fit (P3a:  $\chi^2_{(32)} = 44.42$ ,  $p = 0.071$ ; P3b:  $\chi^2_{(32)} = 45.47$ ,  $p = 0.058$ ).

## Discussion

The current study aimed to investigate whether cerebellar output is required for reinforcement learning-prediction error (RL-PE) coding in the ACC-generated FRN. We studied this in two cerebellar lesion models (cerebellar stroke patients and single-pulse cerebellar TMS) during a probabilistic feedback learning task. While we found RL-PE coding in the FRN for healthy controls and for control stimulation (vertex) in negative outcome/feedback contexts, it was largely absent for cerebellar stroke patients and for cerebellar TMS. The results provide evidence that RL-PE computation is dependent on cerebellar output. Behavioral deficits, however, were subtle. While overall learning success was unaffected by cerebellar lesions or cerebellar TMS, behavioral flexibility, as indexed by choice switching, was reduced. Subtle deficits may be due to compensation by other brain areas within the reinforcement learning network.



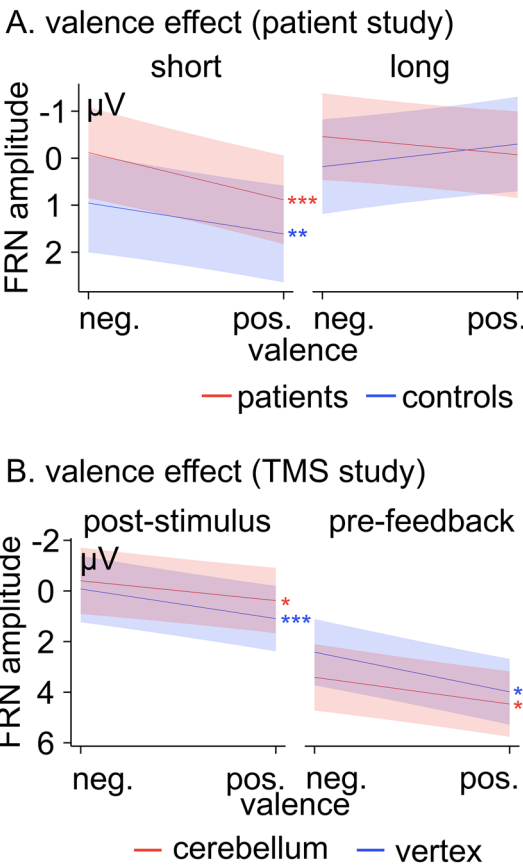
**Table 9.** Inferential statistics for the LME analysis examining the effect of unsigned RL-PE, feedback valence, stimulation site, and TMS timing onto FRN amplitude in Experiment 2 (TMS study)

Fixed effects					
	Est/ $\beta$	SE	df	<i>t</i>	<i>p</i>
(Intercept)	1.92	0.64	21.08	3.02	0.007
Unsigned PE	0.20	0.20	7,125.07	1.00	0.320
Feedback valence	1.08	0.14	7,116.94	7.90	<0.001
Stimulation site	0.08	0.14	5,953.65	0.56	0.577
TMS timing	3.38	0.15	5,521.83	22.99	<0.001
Unsigned PE $\times$ feedback valence	−1.11	0.49	6,506.98	−2.28	0.023
Unsigned PE $\times$ stimulation site	−0.46	0.40	7,120.14	−1.17	0.244
Feedback valence $\times$ stimulation site	−0.27	0.27	7,123.63	−1.00	0.317
Unsigned PE $\times$ TMS timing	0.89	0.40	7,121.23	2.24	0.025
Feedback valence $\times$ TMS timing	0.49	0.27	7,122.51	1.80	0.071
Stimulation site $\times$ TMS timing	1.28	0.31	3,279.84	4.14	<0.001
Unsigned PE $\times$ feedback valence $\times$ stimulation site	2.82	0.93	5,523.19	3.05	0.002
Unsigned PE $\times$ feedback valence $\times$ TMS timing	2.45	0.89	6,049.33	2.75	0.006
Unsigned PE $\times$ stimulation site $\times$ TMS timing	0.38	0.80	7,124.94	0.48	0.632
Feedback valence $\times$ stimulation site $\times$ TMS timing	0.21	0.54	7,124.02	0.39	0.696
Unsigned PE $\times$ feedback valence $\times$ stimulation site $\times$ TMS timing	5.31	4.07	19.75	1.30	0.207
Random effects					
	Variance	SD	Corr		
Subject (intercept)	8.82	2.97			
Subject (unsigned PE $\times$ feedback valence $\times$ stimulation site $\times$ TMS timing)	287.19	16.95	0.03		
Residual	29.10	5.39			
Model fit					
	Marginal		Conditional		
$R^2$	0.08		0.30		

Key:  $p$  values for fixed effects calculated using Satterthwaite's approximations. Model equation:  $FRN \sim 1 + \text{unsigned PE} * \text{feedback valence} * \text{stimulation site} * \text{TMS timing} + (1 + \text{unsigned PE} * \text{feedback valence} * \text{stimulation site} * \text{TMS timing} | \text{subject})$ .  $n_{\text{subjects}} = 22$ ,  $n_{\text{observations}} = 7,162$ .

For healthy controls (Experiment 1) and control stimulation (Experiment 2), RL-PE coding in the FRN was found only for negative feedback. This is consistent with previous studies in healthy participants that found the RL-PE reflected in the FRN only, or at least more strongly, for negative outcomes/feedback (Hoy et al., 2021; Rawls and Lamm, 2021). Note that the direction of the effect was unexpectedly reversed for control stimulation in Experiment 2, as discussed below. Importantly, patients with cerebellar damage and healthy participants receiving cerebellar TMS showed no significant RL-PE coding in the FRN. Activity consistent with RL-PE has previously been described in the rodent cerebellum, although mainly for reward contexts (Kostadinov and Häusser, 2022). It is thus conceivable that RL-PE processing in the cerebellum, as found in these previous studies, is necessary for further RL-PE processing in connected areas classically associated with reinforcement learning (i.e., forebrain and midbrain areas). Of note, feedback processing was not deficient in general, with an intact valence effect in the FRN under short feedback delays in cerebellar patients and in healthy participants receiving cerebellar TMS.

Closer inspection of Figure 7A reveals that the lack of differentiation between low and high RL-PEs in the FRN in patients



**Figure 12.** Slope estimates for FRN amplitude predicted by feedback valence. **A**, Slope estimates for FRN amplitude predicted by feedback valence and modulated by feedback delay and group. Red lines denote patients and blue lines healthy controls. **B**, Slope estimates for FRN amplitude predicted by feedback valence and modulated by TMS timing and stimulation site. Red lines denote cerebellar TMS and blue lines vertex TMS. Colored bands indicate 95% confidence intervals. \* $p < 0.05$ . \*\* $p < 0.01$ . \*\*\* $p < 0.001$ .

was driven by an increase in FRN amplitudes for low RL-PEs rather than a decrease in FRN amplitudes for high RL-PEs. Thus, it appears that the effect is driven by over-activation toward expected outcomes rather than underactivation for unexpected outcomes, which may be indicative of exaggerated perceived salience of expected feedback.

The deficit in RL-PE coding in the FRN was most pronounced in patients with lesions at the border of Crus II and lobule VIIb and to a smaller degree in medial Crus II, Crus I, lobule VIIb, and VIIa. Especially Crus I and II have previously been identified to be involved in decision making and executive control (Berlijn et al., 2024b) and are also connected to the reinforcement learning network (Habas, 2021). The present cluster in Crus II/lobule VIIb overlapped with regions associated with higher cognitive functions, in particular working memory (region D2 in Nettekoven et al., 2024), shown in cerebellar parcellations based on functional magnetic resonance imaging (fMRI) data. The cluster in Crus I, Crus II, and lobule VIIb/VIIa seemed to be related to default mode network/theory of mind, working memory, and spatial rotation/simulation (regions S3, D1, and A1 in Nettekoven et al., 2024). Data on reinforcement learning tasks, however, were not included in the fMRI data on which these parcellations are based. Notably, the cluster in Crus II/lobule VIIb in the present study seems to match with a cluster found in relation to behavioral changes in a feedback learning task very similar to ours (Peterburs et al., 2018). While the abovementioned regions

**Table 10. Model comparisons for the choice switching LME analysis with FRN, P3a, and P3b as additional predictors, respectively, for Experiments 1 and 2**

Model	$n_{\text{parameter}}$	AIC	BIC	Log likelihood	Deviance	Model comparison		
						$\chi^2$	df	$p$
Experiment 1								
Standard model (data with valid FRN, P3a, and P3b)	36	42,815	43,091	−21,372	42,743			
Model with FRN as an additional predictor	68	42,843	43,365	−21,354	42,707	36.01	32	0.286
Model with P3a as an additional predictor	68	42,821	43,343	−21,343	42,685	58.19	32	0.003
Model with P3b as an additional predictor	68	42,825	43,346	−21,344	42,689	54.72	32	0.007
Experiment 2								
Standard model (data with valid FRN)	36	17,686	17,930	−8,807.1	17,614			
Model with FRN as an additional predictor	68	17,725	18,186	−8,794.7	17,589	24.82	32	0.813
Standard model (data with valid P3a/b)	36	17,821	18,065	−8,874.4	17,749			
Model with P3a as an additional predictor	68	17,841	18,302	−8,852.2	17,705	44.42	32	0.071
Model with P3b as an additional predictor	68	17,839	18,301	−8,851.7	17,703	45.47	32	0.058

Model comparisons based on deviance.

were associated with especially aberrant RL-PE coding, the distribution of difference FRN was overall shifted for patients compared with controls (Fig. 8C). Considering the overall lesion distribution in patients, it is conceivable that other posterolateral cerebellar regions also play a role in RL-PE processing.

It has to be noted that behavioral changes associated with cerebellar lesions or cerebellar TMS were quite subtle which seems unexpected given the substantial changes in the FRN. While the FRN generator, ACC, is essential for action–outcome learning (Rudebeck et al., 2008; Camille et al., 2011), the present experiments used additional visual stimuli to represent choices (i.e., button presses). It is therefore conceivable that areas involved in stimulus–outcome learning, such as the orbitofrontal cortex (OFC; Rudebeck et al., 2008; Camille et al., 2011), were able to compensate for ACC-driven deficits. Action–outcome– and stimulus–outcome–based learning might have redundancy to accommodate different environmental requirements. While both the ACC and the OFC receive a wide range of input (Heilbronner and Hayden, 2016; Groman et al., 2021), they might differ in relation to cerebellar input and dependency on predictive information processed by the cerebellum (Peterburs and Desmond, 2016). Previous studies in rodents showed that the cerebellum modulates dopaminergic activity in the substantia nigra (Washburn et al., 2024), and projections from the cerebellum to the VTA were able to modulate place preference (Carta et al., 2019). Both the substantia nigra and the VTA project toward the ACC (Zhang et al., 2017; Elston et al., 2018, 2019). The OFC, in turn, may be more independent from cerebellar processing. While there is some evidence for connections between the cerebellum and OFC (Palesi et al., 2017), we did not measure proxies of OFC activity and therefore cannot conclude whether processing in the OFC was affected. Notably, there seems to be a general pattern of reduced behavioral flexibility and intact acquisition concomitant with cerebellar damage/disruption: learning acquisition was shown to be intact in patients with cerebellar stroke (Thoma et al., 2008; Rustemeier et al., 2016; present Exp. 1) and cerebellar degeneration (A.M. Berlijn, D.M. Huvermann, E. Bechler et al., unpublished observation) as well as healthy participants receiving cerebellar single-pulse or rTMS (present Exp. 2, Kruithof et al., 2025). Nicholas et al. (2024) showed deficits in behavioral flexibility in patients with cerebellar degeneration using a task with constantly changing drifting reward probabilities. In an additional exploratory analysis, we were able to link deficits not in the FRN but in later feedback processing (P3a) to deficits in choice switching, which is consistent with the current conception that the P3 has

stronger behavioral correlates than the FRN (Ullsperger, 2024). There is some evidence for a relation between FRN and behavioral adjustment (Fischer and Ullsperger, 2013; Kirschner et al., 2022; but also see Chase et al., 2011), which might not have played a big enough role in the current study. FRN seems to predict choice switching in highly adaptive environments (Cohen and Ranganath, 2007). The FRN might thus be a readout of a RL-PE that depends on cerebellar output but is not strictly required for learning success in tasks which do not require a high degree of behavioral flexibility. Thoma et al. (2008) and Kruithof et al. (2025) showed that behavioral flexibility required in reversal learning is indeed impaired in cerebellar damage/disruption even in the presence of intact learning acquisition.

In Experiment 2, we varied TMS pulse timing to test whether the cerebellum is potentially involved in response and/or feedback processing selectively. The variation in pulse timing did not appear to modulate the effect of cerebellar TMS on RL-PE coding in the FRN. While this might be related to differential contributions of predictive effects (poststimulus TMS) and feedback processing (prefeedback TMS), an absence of a timing effect might also result from nonoptimal temporal placement of stimulation timings.

While in Experiment 1, the FRN reflected the RL-PE in the expected direction in control participants (i.e., more negative FRN amplitudes for higher RL-PEs), the direction was unexpectedly reversed for the control stimulation in Experiment 2. It thus seems that even though vertex is a common site for control stimulation, it did have an effect on feedback processing, challenging its use as a control condition in Experiment 2. At least one study (Jung et al., 2016) showed reduced activity in the ACC (i.e., the generator of FRN; Hauser et al., 2014) with vertex TMS, although not significantly with the inverted stimulation that we used. Nevertheless, it is still noteworthy that findings for the cerebellar TMS in Experiment 2 replicated the findings in Experiment 1 for cerebellar stroke patients. Both showed a lack of RL-PE coding in FRN. Even though the RL-PE was reflected in the FRN for vertex TMS in the opposite direction, there was a significant RL-PE coding for vertex TMS, while it could not be found for the cerebellar TMS, as well as no effect of stimulation site on learning, thus replicating the overall result pattern in Experiment 1.

In summary, feedback processing, as indexed by the FRN, was shown to be dependent on cerebellar output. While cerebellar dysfunction or damage resulted in only subtle changes in behavioral flexibility with reinforcement learning performance largely intact, processing of RL-PEs as reflected in the FRN was substantially blunted. Crucially, this pattern was consistent across two

**Table 11. Inferential statistics for the LME analysis examining the effect of P3a, feedback valence, response type, feedback delay, group, and block onto choice switching in Experiment 1 (patient study)**

Fixed effects					
	Est/ $\beta$	SE	df	<i>t</i>	<i>p</i>
(Intercept)	0.09	0.04	64.23	2.31	0.024
P3a	0.00	0.00	15,752.83	−1.38	0.167
Feedback valence	−0.18	0.03	15,706.55	−5.87	<0.001
Response type	−0.25	0.03	15,739.13	−8.04	<0.001
Feedback delay	−0.09	0.03	15,714.33	−2.78	0.005
Group	−0.04	0.08	64.23	−0.48	0.632
Block	−0.08	0.02	15,707.24	−5.59	<0.001
P3a × feedback valence	0.00	0.00	15,709.07	−1.06	0.288
P3a × response type	0.01	0.00	15,713.92	1.80	0.072
Feedback valence × response type	−0.13	0.06	15,704.73	−2.14	0.032
P3a × feedback delay	0.00	0.00	15,729.44	−0.82	0.412
Feedback valence × feedback delay	0.12	0.06	15,707.21	1.93	0.053
Response type × feedback delay	−0.10	0.06	15,717.62	−1.61	0.108
P3a × group	0.01	0.00	15,752.83	2.04	0.042
Feedback type × group	−0.08	0.06	15,706.55	−1.34	0.180
Response type × group	−0.07	0.06	15,739.13	−1.09	0.275
Feedback delay × group	0.10	0.06	15,714.33	1.68	0.092
P3a × block	0.00	0.00	15,710.76	1.32	0.186
Feedback valence × block	0.01	0.03	15,706.32	0.33	0.738
Response type × block	−0.10	0.03	15,715.07	−3.29	0.001
Feedback delay × block	0.06	0.03	15,704.81	2.12	0.034
Group × block	0.01	0.03	15,707.24	0.39	0.700
P3a × feedback valence × response type	0.02	0.01	15,707.98	1.78	0.075
P3a × feedback valence × feedback delay	−0.01	0.01	15,708.47	−0.67	0.505
P3a × response type × feedback delay	0.00	0.01	15,712.02	−0.45	0.652
Feedback valence × response type × feedback delay	−0.07	0.12	15,704.17	−0.58	0.565
P3a × feedback valence × group	0.01	0.01	15,709.07	1.18	0.240
P3a × response type × group	0.00	0.01	15,713.92	−0.06	0.951
Feedback valence × response type × group	0.23	0.12	15,704.73	1.87	0.061
P3a × feedback delay × group	−0.02	0.01	15,729.44	−2.23	0.026
Feedback valence × feedback delay × group	0.26	0.12	15,707.21	2.14	0.033
Response type × feedback delay × group	0.03	0.12	15,717.62	0.22	0.829
P3a × feedback valence × block	0.00	0.00	15,706.65	0.33	0.740
P3a × response type × block	0.00	0.00	15,708.16	0.18	0.857
Feedback valence × response type × block	0.09	0.06	15,704.32	1.58	0.114
P3a × feedback delay × block	−0.01	0.00	15,707.03	−1.29	0.199
Feedback valence × feedback delay × block	0.01	0.06	15,705.59	0.21	0.836
Response type × feedback delay × block	−0.06	0.06	15,708.86	−0.92	0.357
P3a × group × block	0.01	0.00	15,710.76	2.05	0.041
Feedback valence × group × block	−0.10	0.06	15,706.32	−1.69	0.092
Response type × group × block	0.00	0.06	15,715.07	0.02	0.983
Feedback delay × group × block	0.06	0.06	15,704.81	1.05	0.295
P3a × feedback valence × response type × feedback delay	0.02	0.02	15,705.28	0.89	0.374
P3a × feedback valence × response type × group	−0.03	0.02	15,707.98	−1.85	0.065
P3a × feedback valence × feedback delay × group	0.02	0.02	15,708.47	1.23	0.219
P3a × response type × feedback delay × group	−0.04	0.02	15,712.02	−2.37	0.018
Feedback valence × response type × feedback delay × group	−0.43	0.25	15,704.17	−1.75	0.081
P3a × feedback valence × response type × block	−0.01	0.01	15,706.99	−1.45	0.147
P3a × feedback valence × feedback delay × block	−0.01	0.01	15,706.18	−1.11	0.268
P3a × response type × feedback delay × block	0.01	0.01	15,707.24	1.29	0.196
Feedback valence × response type × feedback delay × block	−0.14	0.12	15,704.08	−1.13	0.257
P3a × feedback valence × group × block	0.02	0.01	15,706.65	2.31	0.021
P3a × response type × group × block	−0.01	0.01	15,708.16	−1.51	0.132
Feedback valence × response type × group × block	−0.06	0.12	15,704.32	−0.47	0.638
P3a × feedback delay × group × block	−0.01	0.01	15,707.03	−1.13	0.257
Feedback valence × feedback delay × group × block	−0.14	0.12	15,705.59	−1.16	0.246
Response type × feedback delay × group × block	0.07	0.12	15,708.86	0.62	0.533
P3a × feedback valence × response type × feedback delay × group	0.08	0.04	15,705.28	2.16	0.031
P3a × feedback valence × response type × feedback delay × block	0.00	0.02	15,705.60	−0.13	0.894
P3a × feedback valence × response type × group × block	−0.03	0.02	15,706.99	−1.48	0.140
P3a × feedback valence × feedback delay × group × block	−0.01	0.02	15,706.18	−0.36	0.720
P3a × response type × feedback delay × group × block	0.02	0.02	15,707.24	0.93	0.350
Feedback valence × response type × feedback delay × group × block	−0.04	0.24	15,704.08	−0.15	0.883
P3a × feedback valence × response type × feedback delay × group × block	−0.02	0.04	15,705.60	−0.52	0.606

(Table continues.)



**Table 11. Continued**

Fixed effects					
	Est/ $\beta$	SE	df	<i>t</i>	<i>p</i>
Random effects					
	Variance	SD			
Subject (intercept)	0.07	0.26			
Residual	0.86	0.93			
Model fit					
<i>R</i> <sup>2</sup>	Marginal 0.06		Conditional 0.13		
Simple slope analyses					
P3a $\times$ group	Est/ $\beta$	SE		<i>t</i>	<i>p</i>
P3a slope for controls	−0.01	0.00		2.45	0.028
P3a slope for patients	0.00	0.00		0.90	0.734
P3a $\times$ group $\times$ block	Est/ $\beta$	SE		<i>t</i>	<i>p</i>
P3a slope for controls and early experiment ( <i>M</i> − 1SD)	−0.01	0.00		1.56	0.471
P3a slope for controls and late experiment ( <i>M</i> + 1SD)	−0.01	0.00		1.98	0.191
P3a slope for patients and early experiment ( <i>M</i> − 1SD)	0.00	0.00		1.09	>0.999
P3a slope for patients and late experiment ( <i>M</i> + 1SD)	0.01	0.00		2.38	0.068
P3a $\times$ group $\times$ block $\times$ feedback valence	Est/ $\beta$	SE		<i>t</i>	<i>p</i>
P3a slope for controls, early experiment ( <i>M</i> − 1SD), and negative feedback	−0.01	0.01		1.20	>0.999
P3a slope for controls, early experiment ( <i>M</i> − 1SD), and positive feedback	0.00	0.00		1.05	>0.999
P3a slope for controls, late experiment ( <i>M</i> + 1SD), and negative feedback	0.00	0.01		0.04	>0.999
P3a slope for controls, late experiment ( <i>M</i> + 1SD), and positive feedback	−0.02	0.01		2.31	0.168
P3a slope for patients, early experiment ( <i>M</i> − 1SD), and negative feedback	0.00	0.01		0.01	>0.999
P3a slope for patients, early experiment ( <i>M</i> − 1SD), and positive feedback	−0.01	0.00		1.57	>0.999
P3a slope for patients, late experiment ( <i>M</i> + 1SD), and negative feedback	0.01	0.01		0.76	>0.999
P3a slope for patients, late experiment ( <i>M</i> + 1SD), and positive feedback	0.01	0.01		2.55	0.087
P3a $\times$ group $\times$ feedback delay	Est/ $\beta$	SE		<i>t</i>	<i>p</i>
P3a slope for controls and short feedback delays	−0.01	0.00		2.39	0.067
P3a slope for controls and long feedback delays	0.00	0.00		1.06	>0.999
P3a slope for patients and short feedback delays	0.01	0.01		2.15	0.125
P3a slope for patients and long feedback delays	−0.01	0.00		1.44	0.600
P3a $\times$ group $\times$ feedback delay $\times$ response type	Est/ $\beta$	SE		<i>t</i>	<i>p</i>
P3a slope for controls, short feedback delays, and incorrect choices	−0.01	0.01		1.49	>0.999
P3a slope for controls, short feedback delays, and correct choices	−0.01	0.00		1.96	0.403
P3a slope for controls, long feedback delays, and incorrect choices	−0.02	0.01		2.08	0.301
P3a slope for controls, long feedback delays, and correct choices	0.00	0.00		0.72	>0.999
P3a slope for patients, short feedback delays, and incorrect choices	0.00	0.01		0.03	>0.999
P3a slope for patients, short feedback delays, and correct choices	0.02	0.00		3.48	0.004
P3a slope for patients, long feedback delays, and incorrect choices	0.00	0.01		0.72	>0.999
P3a slope for patients, long feedback delays, and correct choices	−0.01	0.00		1.28	>0.999
P3a $\times$ group $\times$ feedback delay $\times$ response type $\times$ feedback valence	Est/ $\beta$	SE		<i>t</i>	<i>p</i>
P3a slope for controls, short feedback delays, incorrect choices, and negative feedback	0.00	0.01		0.07	>0.999
P3a slope for controls, short feedback delays, incorrect choices, and positive feedback	−0.02	0.02		1.53	>0.999
P3a slope for controls, short feedback delays, correct choices, and negative feedback	−0.02	0.01		1.76	>0.999
P3a slope for controls, short feedback delays, correct choices, and positive feedback	0.00	0.00		0.41	>0.999
P3a slope for controls, long feedback delays, incorrect choices, and negative feedback	0.00	0.00		0.64	>0.999
P3a slope for controls, long feedback delays, incorrect choices, and positive feedback	−0.03	0.01		2.41	0.254
P3a slope for controls, long feedback delays, correct choices, and negative feedback	0.00	0.01		0.51	>0.999
P3a slope for controls, long feedback delays, correct choices, and positive feedback	0.00	0.00		0.77	>0.999
P3a slope for patients, short feedback delays, incorrect choices, and negative feedback	−0.01	0.01		1.21	>0.999
P3a slope for patients, short feedback delays, incorrect choices, and positive feedback	0.01	0.02		0.50	>0.999
P3a slope for patients, short feedback delays, correct choices, and negative feedback	0.02	0.01		2.10	>0.999
P3a slope for patients, short feedback delays, correct choices, and positive feedback	0.01	0.00		2.77	0.090
P3a slope for patients, long feedback delays, incorrect choices, and negative feedback	0.00	0.00		0.33	>0.999
P3a slope for patients, long feedback delays, incorrect choices, and positive feedback	−0.01	0.01		1.22	>0.999
P3a slope for patients, long feedback delays, correct choices, and negative feedback	−0.02	0.01		1.46	>0.999
P3a slope for patients, long feedback delays, correct choices, and positive feedback	0.00	0.00		0.03	>0.999

Key: *p* values for fixed effects calculated using Satterthwaite's approximations. Model equation: choice switching  $\sim 1 + P3a * \text{feedback valence} * \text{response type} * \text{feedback delay} * \text{group} * \text{block} + (1 | \text{subject})$ .  $n_{\text{subjects}} = 52$ ,

$n_{\text{observations}} = 15,817$ .

complementary lesion models (i.e., stroke patients and single-pulse TMS). Furthermore, lesion symptom mapping in patients showed that regions at the border of Crus II and Lobule VIIb, medial Crus II, and Crus I were of particular importance.

## References

- Balsters JH, Apps MAJ, Bolis D, Lehner R, Gallagher L, Wenderoth N (2017) Disrupted prediction errors index social deficits in autism spectrum disorder. *Brain* 140:235–246.
- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67:1–48.
- Beck AT, Epstein N, Brown G, Steer RA (1988) An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol* 56:893–897.
- Beck AT, Steer RA, Brown G (1996) *Manual for the Beck depression inventory-II*. San Antonio, TX: Psychological Corporation.
- Berlijn AM, Huvermann DM, Groiss SJ, Schnitzler A, Mittelstaedt M, Bellebaum C, Timmann D, Minnerop M, Peterburs J (2024a) The effect of cerebellar TMS on error processing: a combined single-pulse TMS and ERP study. *Imaging Neurosci* 2:1–19.
- Berlijn AM, Huvermann DM, Schneider S, Bellebaum C, Timmann D, Minnerop M, Peterburs J (2024b) The role of the human cerebellum for learning from and processing of external feedback in non-motor learning: a systematic review. *Cerebellum* 23:1532–1551.
- Bertazzoli G, Esposito R, Mutanen TP, Ferrari C, Ilmoniemi RJ, Miniussi C, Bortoletto M (2021) The impact of artifact removal approaches on TMS-EEG signal. *Neuroimage* 239:118272.
- Brauer M, Curtin JJ (2018) Linear mixed-effects models and the analysis of nonindependent data: a unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychol Methods* 23:389–411.
- Brett M, Leff AP, Rorden C, Ashburner J (2001) Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage* 14:486–500.
- Brunner E, Munzel U (2000) The nonparametric Behrens-Fisher problem: asymptotic theory and a small-sample approximation. *Biom J* 42:17–25.
- Camille N, Tsuchida A, Fellows LK (2011) Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *J Neurosci* 31:15048.
- Carta I, Chen CH, Schott AL, Dorizan S, Khodakhah K (2019) Cerebellar modulation of the reward circuitry and social behavior. *Science* 363:eaav0581.
- Chase HW, Swainson R, Durham L, Benham L, Cools R (2011) Feedback-related negativity codes prediction error but not behavioral adjustment during probabilistic reversal learning. *J Cogn Neurosci* 23:936–946.
- Cohen MX, Ranganath C (2007) Reinforcement learning signals predict future decisions. *J Neurosci* 27:371.
- Cook RD (1977) Detection of influential observation in linear regression. *Technometrics* 19:15–18.
- Deistung A, et al. (2022) Quantitative susceptibility mapping reveals alterations of dentate nuclei in common types of degenerative cerebellar ataxias. *Brain Commun* 4:fcab306.
- Deistung A, Stefanescu MR, Ernst TM, Schlamann M, Ladd ME, Reichenbach JR, Timmann D (2016) Structural and functional magnetic resonance imaging of the cerebellum: considerations for assessing cerebellar ataxias. *Cerebellum* 15:21–25.
- Delorme A, Makeig S (2004) EEGLAB: an open-source toolbox for analysis of single-trial EEG dynamics. *J Neurosci Methods* 134:9–21.
- Diedrichsen J (2006) A spatially unbiased atlas template of the human cerebellum. *Neuroimage* 33:127–138.
- Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N (2009) A probabilistic MR atlas of the human cerebellum. *Neuroimage* 46:39–46.
- Diedrichsen J, Maderwald S, Küper M, Thürling M, Rabe K, Gizewski ER, Ladd ME, Timmann D (2011) Imaging the deep cerebellar nuclei: a probabilistic atlas and normalization procedure. *Neuroimage* 54:1786–1794.
- Diedrichsen J, Zotow E (2015) Surface-based display of volume-averaged cerebellar imaging data. *PLoS One* 10:e0133402.
- Duecker F, Sack AT (2015) Rethinking the role of sham TMS. *Front Psychol* 6:210.
- Elston TW, Croy E, Bilkey DK (2019) Communication between the anterior cingulate cortex and ventral tegmental area during a cost-benefit reversal task. *Cell Rep* 26:2353–2361.e3.
- Elston TW, Kalhan S, Bilkey DK (2018) Conflict and adaptation signals in the anterior cingulate cortex and ventral tegmental area. *Sci Rep* 8:11732.
- Eppinger B, Kray J, Mock B, Mecklinger A (2008) Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia* 46:521–539.
- Filley CM, Fields RD (2016) White matter and cognition: making the connection. *J Neurophysiol* 116:2093–2104.
- Fischer AG, Ullsperger M (2013) Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron* 79:1243–1255.
- Foerde K, Shohamy D (2011) Feedback timing modulates brain systems for learning in humans. *J Neurosci* 31:13157.
- Foti D, Weinberg A, Bernat EM, Proudfit GH (2015) Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. *Clin Neurophysiol* 126:1338–1347.
- Fouragnan E, Retzler C, Philastides MG (2018) Separate neural representations of prediction error valence and surprise: evidence from an fMRI meta-analysis. *Hum Brain Mapp* 39:2887–2906.
- Gatti D, Rinaldi L, Vecchi T, Ferrari C (2023) Understanding cerebellar cognitive and social functions: methodological challenges and new directions for future transcranial magnetic stimulation studies. *Curr Opin Behav Sci* 53:101300.
- Groman SM, Lee D, Taylor JR (2021) Unlocking the reinforcement-learning circuits of the orbitofrontal cortex. *Behav Neurosci* 135:120.
- Habas C (2021) Functional connectivity of the cognitive cerebellum. *Front Syst Neurosci* 15:642225.
- Halahakoon DC, Kieslich K, O'Driscoll C, Nair A, Lewis G, Roiser JP (2020) Reward-processing behavior in depressed participants relative to healthy volunteers: a systematic review and meta-analysis. *JAMA Psychiatry* 77:1286–1295.
- Hardwick RM, Lesage E, Miall RC (2014) Cerebellar transcranial magnetic stimulation: the role of coil geometry and tissue depth. *Brain Stimul* 7:643–649.
- Hauser TU, Iannaccone R, Stämpfli P, Drechsler R, Brandeis D, Walitza S, Brem S (2014) The feedback-related negativity (FRN) revisited: new insights into the localization, meaning and network organization. *Neuroimage* 84:159–168.
- Heilbronner SR, Hayden BY (2016) Dorsal anterior cingulate cortex: a bottom-up view. *Annu Rev Neurosci* 39:149–170.
- Heise MJ, Mon SK, Bowman LC (2022) Utility of linear mixed effects models for event-related potential research with infants and children. *Dev Cogn Neurosci* 54:101070.
- Hinneberg BM, Hegele M (2022) Acting in temporal contexts: on the behavioral and neurophysiological consequences of feedback delays. *Neuroscience* 486:91–102.
- Holroyd CB, Coles MGH (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 109:679–709.
- Hoy CW, Steiner SC, Knight RT (2021) Single-trial modeling separates multiple overlapping prediction errors during reward processing in human EEG. *Commun Biol* 4:910.
- Hruby T, Marsalek P (2003) Event-related potentials-the P3 wave. *Acta Neurobiol Exp* 63:55–63.
- Ichikawa N, Siegle GJ, Dombrovski A, Ohira H (2010) Subjective and model-estimated reward prediction: association with the feedback-related negativity (FRN) and reward prediction error in a reinforcement learning task. *Int J Psychophysiol* 78:273–283.
- Jung J, Bungert A, Bowtell R, Jackson SR (2016) Vertex stimulation as a control site for transcranial magnetic stimulation: a concurrent TMS/fMRI study. *Brain Stimul* 9:58–64.
- Kathagen T, Kaminski J, Heinz A, Buchert R, Schlagenhaut F (2020) Striatal dopamine and reward prediction error signaling in unmedicated schizophrenia patients. *Schizophr Bull* 46:1535–1546.
- Keren H, et al. (2018) Reward processing in depression: a conceptual and meta-analytic review across fMRI and EEG studies. *Am J Psychiatry* 175:1111–1120.
- Kirschner H, Fischer AG, Ullsperger M (2022) Feedback-related EEG dynamics separately reflect decision parameters, biases, and future choices. *Neuroimage* 259:119437.
- Kostadinov D, Häusser M (2022) Reward signals in the cerebellum: origins, targets, and functional implications. *Neuron* 110:1290–1303.
- Kruihof ES, Drop EM, Gerits D, Klaus J, Schutter DJLG (2025) Continuous theta burst stimulation to the medial posterior cerebellum impairs reversal learning in healthy volunteers. *Cogn Affect Behav Neurosci*.

- Kuznetsova A, Brockhoff P, Christensen R (2017) LmerTest package: tests in linear mixed effects models. *J Stat Softw* 82:1–26.
- Lenth RV (2025) emmeans: Estimated marginal means, aka least-squares means.
- Long JA (2019) interactions: comprehensive, user-friendly toolkit for probing interactions. Available at: <https://cran.r-project.org/package=interactions>
- Luke SG (2017) Evaluating significance in linear mixed-effects models in R. *Behav Res Methods* 49:1494–1502.
- Manto M-U, et al. (2024) Consensus paper: cerebellum and reward. *Cerebellum* 23:2169–2192.
- McDougle SD, Butcher PA, Parvin DE, Mushtaq F, Niv Y, Ivry RB, Taylor JA (2019) Neural signatures of prediction errors in a decision-making task are modulated by action execution failures. *Curr Biol* 29:1606–1613.e5.
- Meadows CC, Gable PA, Lohse KR, Miller MW (2016) The effects of reward magnitude on reward processing: an averaged and single trial event-related potential study. *Biol Psychol* 118:154–160.
- Merz J, Lehl S, Galster V, Erzigkeit H (1975) MWT-B - ein intelligenzkurztest. *Psychiatr Neurol Med Psychol* 27:423–428.
- Miltner WHR, Braun CH, Coles MGH (1997) Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a “generic” neural system for error detection. *J Cogn Neurosci* 9:788–798.
- Nettekoven C, Zhi D, Shahshahani L, Pinho AL, Saadon-Grosman N, Buckner RL, Diedrichsen J (2024) A hierarchical atlas of the human cerebellum for functional precision mapping. *Nat Commun* 15:8376.
- Nicholas J, Amlang C, Lin C-YR, Montaser-Kouhsari L, Desai N, Pan M-K, Kuo S-H, Shohamy D (2024) The role of the cerebellum in learning to predict reward: evidence from cerebellar ataxia. *Cerebellum* 23:1355–1368.
- Nieuwenhuis R, te Grotenhuis M, Pelzer B (2012) InfluenceME: tools for detecting influential data in mixed effects models. *R J* 4:38–47.
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97–113.
- Palesi F, De Rinaldis A, Castellazzi G, Calamante F, Muhler N, Chard D, Tournier JD, Magenes G, D’Angelo E, Gandini Wheeler-Kingshott CAM (2017) Contralateral cortico-ponto-cerebellar pathways reconstruction in humans in vivo: implications for reciprocal cerebro-cerebellar structural connectivity in motor and non-motor areas. *Sci Rep* 7:12841.
- Palesi F, Ferrante M, Gaviraghi M, Misiti A, Savini G, Lascialfari A, D’Angelo E, Gandini Wheeler-Kingshott CAM (2021) Motor and higher-order functions topography of the human dentate nuclei identified with tractography and clustering methods. *Hum Brain Mapp* 42:4348–4361.
- Peterburs J, Desmond JE (2016) The role of the human cerebellum in performance monitoring. *Syst Neurosci* 40:38–44.
- Peterburs J, Gajda K, Koch B, Schwarz M, Hoffmann K-P, Daum I, Bellebaum C (2012) Cerebellar lesions alter performance monitoring on the antisaccade task—an event-related potentials study. *Neuropsychologia* 50:379–389.
- Peterburs J, Hofmann D, Becker MPI, Nitsch AM, Miltner WHR, Straube T (2018) The role of the cerebellum for feedback processing and behavioral switching in a reversal-learning task. *Brain Cogn* 125:142–148.
- Peterburs J, Kobza S, Bellebaum C (2016) Feedback delay gradually affects amplitude and valence specificity of the feedback-related negativity (FRN). *Psychophysiology* 53:209–215.
- Phillips JR, Hewedi DH, Eissa AM, Moustafa AA (2015) The cerebellum and psychiatric disorders. *Front Public Health* 3:66.
- Polich J (2007) Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* 118:2128–2148.
- R Core Team (2023) R: a language and environment for statistical computing. Available at: <https://www.R-project.org/>
- Ramnani N (2006) The primate cortico-cerebellar system: anatomy and function. *Nat Rev Neurosci* 7:511–522.
- Rawls E, Lamm C (2021) The aversion positivity: mediofrontal cortical potentials reflect parametric aversive prediction errors and drive behavioral modification following negative reinforcement. *Cortex* 140:26–39.
- Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical conditioning II: current research and theory* (Black AH, Prokasy WF, eds), pp 64–99. New York: Appleton-Century-Crofts.
- Robinson RG, Jorge RE (2016) Post-stroke depression: a review. *Am J Psychiatry* 173:221–231.
- Posit Team (2023) RStudio: Integrated development environment for R. Available at: <http://www.posit.co/>
- Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MFS (2008) Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci* 28:13775.
- Rustemeier M, Koch B, Schwarz M, Bellebaum C (2016) Processing of positive and negative feedback in patients with cerebellar lesions. *Cerebellum* 15:425–438.
- Sambrook TD, Goslin J (2015) A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychol Bull* 141:213.
- San Martín R (2012) Event-related potential studies of outcome processing and feedback-guided learning. *Front Hum Neurosci* 6:304.
- Schmahmann JD, Sherman JC (1998) The cerebellar cognitive affective syndrome. *Brain* 121:561–579.
- Schmitz-Hübsch T, et al. (2006) Scale for the assessment and rating of ataxia. *Neurology* 66:1717.
- Shen N, González B (2021) Bayesian information criterion for linear mixed-effects models. *ArXiv Prepr ArXiv210414725*.
- Sokolov AA, Miall RC, Ivry RB (2017) The cerebellum: adaptive prediction for movement and cognition. *Trends Cogn Sci* 21:313–332.
- Stoodley CJ, MacMore JP, Makris N, Sherman JC, Schmahmann JD (2016) Location of lesion determines motor vs. cognitive consequences in patients with cerebellar stroke. *Neuroimage Clin* 12:765–775.
- Sutton RS, Barto AG (2018) Temporal-difference learning. In: *Reinforcement learning: an introduction*, 2nd ed. pp 119–140. Cambridge, MA: The MIT press.
- Thoma P, Bellebaum C, Koch B, Schwarz M, Daum I (2008) The cerebellum is involved in reward-based reversal learning. *Cerebellum* 7:433–443.
- Timmann D, Küper M, Gizewski ER, Schoch B, Donchin O (2022) Lesion-symptom mapping of the human cerebellum. In: *Handbook of the cerebellum and cerebellar disorders* (Manto M-U, Gruol DL, Schmahmann JD, Koibuchi N, Sillitoe RV, eds), Ed 2, pp 1857–1890. Houston, TX: Springer Cham.
- Ullsperger M (2024) Beyond peaks and troughs: multiplexed performance monitoring signals in the EEG. *Psychophysiology* 61:e14553.
- Volpert-Esmund HI, Page-Gould E, Bartholow BD (2021) Using multilevel models for the analysis of event-related potentials. *Int J Psychophysiol* 162:145–156.
- Wagner AR, Rescorla RA (1972) Inhibition in Pavlovian conditioning: application of a theory. In: *Inhibition and learning* (Boakes RA, Halliday MS, eds), pp 301–336. New York: Academic Press.
- Wahlund LO, et al. (2001) A new rating scale for age-related white matter changes applicable to MRI and CT. *Stroke* 32:1318–1322.
- Washburn S, Oñate M, Yoshida J, Vera J, Bhuvanansundaram R, Khatami L, Nadim F, Khodakhah K (2024) The cerebellum directly modulates the substantia nigra dopaminergic activity. *Nat Neurosci* 27:497–513.
- Wu W, Keller CJ, Rogasch NC, Longwell P, Shpigel E, Rolle CE, Etkin A (2018) ARTIST: a fully automated artifact rejection algorithm for single-pulse TMS-EEG data. *Hum Brain Mapp* 39:1607–1625.
- Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, Gerig G (2006) User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* 31:1116–1128.
- Zhang Y, Larcher KM-H, Misic B, Dagher A (2017) Anatomical and functional organization of the human substantia nigra and its connections. *Johansen-Berg H, ed. Elife* 6:e26653.