Behavioral/Cognitive

The Human Cerebellum Encodes Temporally Sensitive Reinforcement Learning Signals

Juliana E. Trach,[1] Yiran Ou,[1] and Samuel D. McDougle[1,2]
[1]Department of Psychology, Yale University, New Haven, Connecticut 06510 and and [2]Wu Tsai Institute, Yale University, New Haven, Connecticut 06510

In addition to supervised motor learning, the cerebellum also supports nonmotor forms of learning, including reinforcement learning (RL). Recent studies in animal models have identified core RL signals related to reward processing, reward prediction, and prediction errors in specific regions in the cerebellar cortex. However, the constraints on these signals remain poorly understood, particularly in humans. Here, we investigated cerebellar RL signals in a computationally driven fMRI study. Human participants performed an RL task without low-level sensorimotor contingencies ($N = 32$; N female = 24). We observed robust RL signals related to reward processing and reward prediction errors (RPEs) in cognitive regions of the cerebellum. These signals were not explained by oculomotor or physiological confounds. By manipulating the delay between choices and reward outcomes, we discovered that cerebellar RL signals are temporally sensitive: robust when feedback was delivered shortly following choices but undetectable at supra-second feedback delays. Similar delay effects were not found in other areas implicated in reward processing, including the ventral striatum and hippocampus. Furthermore, RPE activity in the cerebellum was related to behavioral performance when feedback was delivered promptly, but not when it was delayed. Connectivity analyses revealed that during RL feedback, cognitive areas of the cerebellum coactivated with a network that included the medial and lateral prefrontal cortex and caudate nucleus. Together, these results highlight a temporally constrained contribution of the human cerebellum to a cognitive learning task.

Key words: cerebellum; human neuroscience; reinforcement learning; reward prediction error

Significance Statement

Reinforcement learning (RL)—the shaping of behavior through reward feedback—is an essential cognitive capacity. Previous work has focused almost exclusively on cerebral circuits that support RL; however, recent work in animal models implicates the cerebellum. Despite growing interest in the "cognitive cerebellum," its contributions to human RL remain unclear. We show that regions spanning lobules Crus I/II respond to rewards and encode reward prediction errors in a temporally sensitive manner. Our results build directly on research in model organisms, highlighting functional parallels across species and the importance of cross talk between human and animal researchers. Including the cerebellum as a node in RL networks creates a potential new target for intervention in cases of reward processing dysfunction, like addiction.

Introduction

A recent paradigm shift in cerebellar neuroscience recognizes that the cerebellum contributes to both motor control and cognition (Leiner et al., 1986; Ito, 2008; Strick et al., 2009; Stoodley et al., 2012; Buckner, 2013; Sokolov et al., 2017; Diedrichsen et al., 2019; Schmahmann, 2019; Diedrichsen and McDougle, 2026). This shift is motivated by converging lines of evidence: The cerebellum is implicated in cognitive neurodevelopmental disorders (Stoodley, 2016), it is activated by nonmotor tasks (Van Overwalle et al., 2014, 2020; King et al., 2019), damage produces cognitive deficits (Schmahmann and Sherman, 1998; Drepper et al., 1999; Timmann et al., 2010; Butcher et al., 2017; McDougle et al., 2022; Nicholas et al., 2023), and it has bidirectional connections with nonmotor areas of the cerebrum (Middleton and Strick, 1994; Kelly and Strick, 2003; Hoshi et al., 2005; Strick et al., 2009; Bostan et al., 2010, 2013; Buckner et al., 2011; Buckner, 2013; Wagner and Luo, 2020; King et al., 2023). Still, a unifying framework of cerebellar function remains elusive. What specific cognitive computations might the cerebellum implement?

Some recent work on this question focuses on cerebellar contributions to reinforcement learning (RL; Hull, 2020; Kostadinov and Häusser, 2022). Studies in rodents and nonhuman primates

Received Dec. 19, 2025; revised May 5, 2026; accepted May 22, 2026.

Author contributions: J.E.T. and S.D.M. designed research; J.E.T. and Y.O. performed research; J.E.T. analyzed data; J.E.T. and S.D.M. wrote the paper.

We thank the ACT lab and the other Cog Neuro labs at Yale for their helpful discussions about this project. These data were collected at BrainWorks at the WuTsai Institute at Yale University. We thank Roeland Hancock and Alex Forrence for their technical support at BrainWorks and Tess Levy, Sanghoon Kang, Laurent Caplette, Lily Behm, Diana Wei, Omri Raccah, and Sophie Allen for their support with data collection. J.E.T. is supported by the NSF GRFP. S.D.M. is supported by NIH Grant Number R01 NS132926.

The authors declare no competing financial interests.

Correspondence should be addressed to Juliana E. Trach at juliana.trach@yale.edu.

This paper contains supplemental material available at: https://doi.org/10.1523/JNEUROSCI.2313-25.2026

https://doi.org/10.1523/JNEUROSCI.2313-25.2026

have revealed cerebellar activity encoding core components of RL, including reward processing, reward expectation, and reward prediction errors (RPEs; Wagner et al., 2017; Heffley et al., 2018; Kostadinov et al., 2019; Sendhilnathan et al., 2020). Such signals have been observed in granule cells (Wagner et al., 2017), climbing fiber inputs (Heffley and Hull, 2019; Kostadinov et al., 2019; Larry et al., 2019), and Purkinje cells (Sendhilnathan et al., 2020) in areas spanning lobules Crus I and Crus II (Heffley and Hull, 2019; Sendhilnathan et al., 2020). Moreover, lesion and stimulation studies point to the causal role of the cerebellum in RL (McDougle et al., 2016; Carta et al., 2019; Nicholas et al., 2023; Huvermann et al., 2025). There is now general consensus that the cerebellum supports RL (Manto et al., 2024) in conjunction with circuits in the cerebrum.

Still, key gaps remain, especially in human neuroscience. First, there are practical issues: full coverage of the cerebellum with human functional neuroimaging is rarely achieved (Wang et al., 2025), and cerebellar BOLD signals are susceptible to motor and physiological confounds (Diedrichsen et al., 2010; Schlerf et al., 2012; Striemer et al., 2015). Second, many RL tasks involve sensorimotor contingencies that blur cognitive and motor components of RL (O'Doherty et al., 2003; Seymour et al., 2004; Sendhilnathan et al., 2020). Finally, specific anatomical localization of RL signals in humans is rarely reported when they are incidentally detected, complicating comparisons to model organisms. These challenges may explain heterogeneity in previously reported cerebellar activations during RL (O'Doherty et al., 2003; Garrison et al., 2013; Kruithof et al., 2023).

Theoretical gaps also remain, particularly regarding constraints on cerebellar RL computations. One candidate constraint is temporal sensitivity: In associative motor learning, the cerebellum is optimized for processing rapid cue–outcome intervals (i.e., <~2 s; Smith, 1968; Ivry and Keele, 1989; Kitazawa et al., 1995; Chettih et al., 2011; Brudner et al., 2016; Avraham et al., 2022), and learning is attenuated when these intervals are exceeded (Gerwig et al., 2008). In contrast, areas like basal ganglia, hippocampus, and neocortex can associate events over longer timescales (Buhusi and Meck, 2005; Foerde and Shohamy, 2011). We thus hypothesize that the cerebellum may be selectively engaged in RL when predictive associations unfold rapidly.

We tested this hypothesis using choice behavior, eyetracking, physiological recordings, and model-based fMRI in humans. Our task avoided low-level sensorimotor associations, and we optimized cerebellar coverage. By manipulating delays between choice and feedback, we tested temporal constraints on cerebellar RL signals. Moreover, we used whole-brain fMRI to compare cerebellar RL signals to those in canonical RL regions, to link neural responses to behavior, and to ask if RL feedback drives functional correlations with other regions of the brain.

Materials and Methods

Participants

Thirty-five individuals participated in the study [mean age = 22.9 (18–33); *N* female = 24]. We planned to exclude participants that were not engaging in the task (not responding to >25% of trials or only responding with one response >90% of trials) or that produced too much head motion during the scan (maximum motion over 2.5 mm). No participants met our behavioral exclusion threshold. Two participants were excluded for excessive head motion, and an additional subject was excluded for technical issues with the data. Our final sample was 32 participants [mean age = 22.8 (18–33); *N* female = 24]. The task protocol was approved by the Yale Institutional Review Board. Participants were compensated $30/h. They were also told that they could receive up to $10 bonus payment for their performance in the task to connect the reward feedback during the task to actual reward. We gave all participants the full bonus payment.

Experimental session

The experimental session was 2 h long. Participants arrived and were consented and screened for metal. In the scanner, they completed three runs of a probabilistic RL task and three runs of a statistical learning (SL) task (not discussed here). The RL task was always performed before the SL task. Each functional run began with an eye-tracker calibration phase (Eyelink 1000 Plus, long-range monocular). After the functional runs, we collected a T1-weighted (T1w) anatomical image before bringing them out of the scanner. Finally, the session concluded with a test phase for the SL task and a debriefing questionnaire. Task and protocol details are explained below.

RL task

Participants performed three runs of a probabilistic RL task in the scanner. During each run, participants saw four pairs of images (eight unique images per run) across trials (Supplemental Fig. 1). On each trial, they saw one pair of images side-by-side and used an MR-safe, two-button button box with their right hand to select one of the two images (index finger to select left image, middle finger to select right image). They received probabilistic reward feedback (reward +1 or nonreward +0) on their response after a delay (Fig. 1*a*). One image in each pair was associated with a 0.25 reward probability and the other with a 0.75 reward probability throughout a run of trials. Participants were instructed to select the stimulus on each trial that they thought was most likely to give them points. They were informed that their point total would be translated into a monetary bonus at the end of the session (up to $10) to motivate them to perform well and ensure that the point-feedback was rewarding. Crucially, the location of each image was counterbalanced across trials so that reward information was associated with the image—the abstract choice—and not with a specific location or motor action.

We manipulated the delay between stimulus offset and the presentation of the reward feedback in order to test the hypothesis that the cerebellum would be particularly involved in processing short-latency feedback signals. Participants viewed the pairs of stimuli for 1.5 s on each trial, during which time they made their response. Once they made their response, the fixation cross changed color to indicate that their choice had been registered and the stimuli remained on the screen for the remainder of the 1.5 s. The stimuli then disappeared, and there was a delay before they were presented with reward feedback for 1 s. To test the hypothesis that the cerebellum would be sensitive to short-latency feedback, we assigned two pairs in each run to a short feedback delay where the feedback appeared 0.8 s after the stimulus disappeared and the other two pairs to a long feedback delay where feedback appeared after 3 s. If they did not respond quickly enough, they received the feedback "Please respond faster." There was a jittered intertrial interval (1–3 s) after feedback presentation.

Prior to the three runs of the task, participants executed a short practice block with two pairs of stimuli to orient them to the task. There was no variation in the delay between responses and feedback in the practice block. During the main runs, participants performed 96 trials during a run (24 presentations per pair randomly interleaved). Novel stimuli were used in each run to prevent participants from using learned values from previous runs.

Behavioral and physiological analysis

We used participant choices and RTs to characterize their behavior. Trials where participants did not respond were excluded from analysis. Learning was operationalized as the probability of selecting the stimulus associated with the higher reward probability across trials. To quantify learning, we implemented a generalized linear model (GLM) that used iteration (how many times the participant had viewed a specific pair of stimuli) to predict whether or not they chose the higher value stimulus (1, chose higher value stimulus; 0, chose lower value stimulus). Overall learning was measured as the probability of selecting the higher value stimulus across all trials. We used *t* tests to compare overall learning and

RTs between short- and long-delay trials and to compare biophysiological measures on rewarded versus nonrewarded trials.

We additionally collected eyetracking data during the entirety of the scan using a long-range Eyelink 1000 plus. We did not obtain eyetracking data from two participants who used MR-safe glasses during the scan which interfered with tracking and an additional participant due to equipment malfunction. We began each run with a calibration phase. We were primarily interested in differences in gaze behavior on rewarded versus nonrewarded trials that might impact our results. To examine this possibility, we summarized gaze behavior within each trial and at feedback by counting the number of saccades in these time intervals and used a *t* test to compare between rewarded and nonrewarded trials.

Finally, we collected pulse and respiration data from each participant using Siemens peripheral devices during scanning (samples every 2.5 ms). We used a pulse-oximeter positioned on the index finger of the participant's left hand to collect pulse data and a respiration belt positioned around the participant's abdomen to track respiration. A portion of the data was lost due to signal drop out during the task or subsequent data storage issues. For five additional participants, we excluded their pulse data from our analyses as the pulse-oximeter was unable to acquire a reliable pulse signal. Thus, of the usable runs of functional BOLD data (86 runs across 32 participants), we retained 74.4% of the respiration data (64 runs across 24 participants) and 55.8% of the pulse data (48 runs across 18 participants) for analysis.

We were primarily interested in whether these physiological metrics varied significantly across rewarded and nonrewarded trials in a way that might impact our results. To test this, we calculated trialwise pulse and respiration rates and then compared these rates on rewarded versus nonrewarded trials. After extracting the data for each metric in each run, we used the *peakdet* function in MATLAB (Billauer, 2009) to locate local minima and maxima in the traces. We then calculated a trialwise rate by counting the number of maxima and minima from the onset of the stimuli on one trial to the onset of the stimuli for the next trial and divided that count by the duration between these two time points (in minutes) to determine the rate. We opted to average the number maxima and minima during the time window of interest in this calculation in order to account for situations where a portion, but not all, of a cardiac or respiratory cycle fell within the window of interest. We then calculated subjectwise averages of these rates for rewarded and nonrewarded trials separately and used a *t* test to compare the rates.

Computational modeling

We conducted an RL modeling analysis on subjects' trial-by-trial behavioral data. These models allowed us to bridge our behavioral and neural data through the lens of RL theory (Sutton and Barto, 1998). Specifically, we aimed to model subject behavior to estimate latent variables that reflect the core prediction error processes of RL and then use these computational estimates to explain variance in neural data using model-based fMRI (Cohen et al., 2017). This computational analysis thus allowed us to generate inferred RPE time courses for use in our neural analyses.

To model RL behavior and RPEs, we employed a well-established RL model with the basic form as follows:

$$Q(s)_{t+1} = Q(s)_t + \alpha\delta, \tag{1}$$

$$\delta = r - Q(s)_t, \tag{2}$$

where the value (Q) of a given stimulus (s) that is chosen by the participant on trial t is updated according to the RPE (δ) on that trial (the difference between the expected value Q and received reward r), with a learning rate parameter α.

We compared the fit of three models to participant data. We fit one variant of the model with a single learning rate:

$$Q(s)_{t+1} = Q(s)_t + \alpha\delta. \tag{3}$$

We fit an additional model with separate learning rates (α^+, α^-) for rewarded versus nonrewarded outcomes (Frank et al., 2007; Trach et al., 2025):

$$Q(s)_{t+1} = \begin{cases} Q(s)_t + \alpha^+\delta & \text{if} \quad r = 1 \\ Q(s)_t + \alpha^-\delta & \text{if} \quad r = 0 \end{cases}. \tag{4}$$

Finally, because our task consisted of stimulus pairs, it is possible that participants may make counterfactual inferences, such that if one stimulus of the pair was rewarding, then the other would be unlikely to give them reward and vice versa (Boorman et al., 2011). In other words, they may update both the Q value of the chosen stimulus, s, and of the unchosen stimulus, s', when they receive feedback about s. We operationalized this by adding two additional free parameters (ω^+, ω^-) that dictated how much feedback about the chosen stimulus would impact the Q value of the unchosen stimulus on rewarded and nonrewarded trials as follows:

$$Q(s)_{t+1} = \begin{cases} Q(s)_t + \alpha^+\delta & \text{if} \quad r = 1 \\ Q(s)_t + \alpha^-\delta & \text{if} \quad r = 0 \end{cases}, \tag{5}$$

$$Q(s')_{t+1} = \begin{cases} Q(s')_t + \omega^+(0 - Q(s')_t) & \text{if} \quad r = 1 \\ Q(s')_t + \omega^-(1 - Q(s')_t) & \text{if} \quad r = 0 \end{cases}. \tag{6}$$

Across all models, action selection between the two presented stimuli was modeled using the softmax function:

$$p(s) = \frac{e^{\beta Q(s)}}{\sum_i e^{\beta Q(s_i)}},$$

where β reflects the inverse softmax temperature. We used the MATLAB function fmincon to fit our model to each subject's observed choice data, optimizing the parameter values to maximize the log posterior probability of the choice data given the model. During fitting, α, (α^+, α^-), and (ω^+, ω^-) were constrained on [0,1] and β on [0,50], and a Gamma (2,3) prior distribution was used to discourage extreme values of β (Leong et al., 2017; Trach et al., 2025). Importantly, we fit one of each fitted parameter per subject rather than fitting new values for each run (i.e., three of each per participant). The fitted parameters are thus optimized to best reflect participant learning and performance across runs. This approach gives us more data to derive a stable estimate of these parameters for each subject. Furthermore, theoretically, these fitted parameters should reflect psychological variables that remain relatively stable within the task context, and fitting runs separately could lead to overfitting. The fitting procedure was conducted 200 times per subject with randomized starting parameter values to avoid local minima. Simulated choice data (from the optimized model) were produced to investigate and visualize the model's ability to replicate the main behavioral findings and to generate the trial-by-trial RPE time courses that were fit to the neural data (see below; see also Supplemental Fig. 2).

We compared models using the Bayesian information criterion (BIC). This fit metric was implemented here as it balances both goodness of fit with model complexity. Smaller values of BIC are indicative of better model fit. We found that the dual learning rate model was the best fit for participant behavior (one learning rate model: summed BIC = 10,291; dual learning rate model: summed BIC = 9,953; inference model: summed BIC = 10,027; Supplemental Fig. 2). Thus, we used the dual learning rate model to simulate RPEs for our analyses, though we note that all models produced similar RPE time courses (Supplemental Fig. 2). Thus, we used the dual learning rate model in our analyses.

fMRI acquisition and preprocessing

MR data were acquired with a 3T Siemens Prisma scanner with a 64-channel head coil at BrainWorks at the WuTsai Institute at Yale University. Whole-brain functional images were collected using an echo-planar imaging sequences with repetition time (TR), 1,000 ms; echo time

(TE), 30 ms; voxel size, $2.5 \times 2.5 \times 3$ mm^3; field of view, $20.8 \times 20.8 \times 20.8$ cm^3; 48 slices; P to A phase encoding direction, with multiband acceleration factor, 3 (interleaved); and in-plane acceleration factor, 2. Gradient echo field maps were acquired to correct for distortions due to B0 inhomogeneities (acquisition parameters, voxel size, $3 \times 3 \times 3$ mm^3; field of view, $24 \times 24 \times 24$ cm^3). We also collected a high-resolution T1w MPRAGE; voxel size, 1 mm^3; and field of view, $25.6 \times 25.6 \times 25.6$ cm^3. As previously noted, we used Siemens peripherals to track pulse and respiration during the functional runs as well.

Anatomical data preprocessing. Preprocessing was performed using fMRIPrep 23.2.1 (Esteban et al., 2018, 2019), which is based on Nipype 1.8.6 (Gorgolewski et al., 2011, 2018). The T1w image was corrected for intensity nonuniformity with N4BiasFieldCorrection (Tustison et al., 2010), distributed with ANTs 2.5.0 (Avants et al., 2008), and used as T1w reference throughout the workflow. The T1w reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white matter (WM), and gray matter (GM) was performed on the brain-extracted T1w using fast (FSL, RRID:SCR_002823; Zhang et al., 2001). Brain surfaces were reconstructed using recon-all (FreeSurfer 7.3.2, RRID:SCR_001847; Dale et al., 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANT-derived and FreeSurfer-derived segmentations of the cortical GM of Mindboggle (RRID:SCR_002438; Klein et al., 2017). Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.5.0), using brain-extracted versions of both T1w reference and the T1w template.

Functional data preprocessing. For each of the BOLD runs per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume was generated, using a custom methodology of fMRIPrep, for use in head-motion correction. Head-motion parameters with respect to the BOLD reference (transformation matrices and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL; Jenkinson et al., 2002). The BOLD reference was then coregistered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration (Greve and Fischl, 2009). Coregistration was configured with 12 df to account for distortions remaining in the BOLD reference. Several confounding time series were calculated based on the preprocessed BOLD: framewise displacement (FD), Derivative of timecourses VARiance over Spatial voxels (DVARS), and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions; Power et al., 2014) and Jenkinson (relative root mean square displacement between affines; Jenkinson et al., 2002). FD and DVARS are calculated for each functional run, both using their implementations in Nipype [following the definitions by Power et al. (2014)]. The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (CompCor; Behzadi et al., 2007). Principal components are estimated after high-pass filtering the preprocessed BOLD time series (using a discrete cosine filter with 128 s cutoff) for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM, and combined CSF + WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. (2007) in that instead of eroding the masks by 2 pixels on BOLD space, a mask of pixels that likely contain a volume fraction of GM is subtracted from the aCompCor masks. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's aseg segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50% of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head-motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardized DVARS were annotated as motion outliers. Additional nuisance time series are calculated by means of principal components analysis of the signal found within a thin band (crown) of voxels around the edge of the brain, as proposed by Patriat et al. (2017). All resamplings can be performed with a single interpolation step by composing all the pertinent transformations (i.e., head-motion transform matrices, susceptibility distortion correction when available, and coregistrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using nitransforms, configured with cubic B-spline interpolation. (Copyright Waiver: Much of the above boilerplate text was automatically generated by fMRIPrep with the express intention that users should copy and paste this text into their manuscripts unchanged. It is released under the CC0 license.)

Additional preprocessing after fMRIprep. The primary aim of this study was to isolate neural correlates of RL within the cerebellum. Thus, we ran planned analyses on masked cerebellar data, in addition to the whole brain. To constrain analyses within the cerebellum and avoid bleed-over from adjacent visual/temporal regions, we masked the cerebellum from functional data aligned to standard MNI space prior to spatial smoothing. Whole-brain and cerebellar-masked data were spatially smoothed with a 5 mm kernel. We excluded runs where participants produced any movement over 2.5 mm (1 voxel; $N = 10$ runs) and excluded participants if they did not have multiple runs to contribute ($N = 2$). One additional participant was excluded due to technical difficulties.

fMRI analysis

Whole-brain and cerebellar univariate analysis. We modeled BOLD responses using GLMs implemented in FSL (version 6.0.7.9). Results from three GLMs are presented in the main text, with additional results from another GLM in the supplement. First, we used a GLM to examine responses to reward versus nonreward feedback. We modeled events as boxcar regressors that lasted for the duration of the event of interest. Regressors of interest were convolved with a double gamma HRF using FSL. Here, we included separate task regressors, in addition to our confound regressors, for (1) choice stimulus appearance, (2) participant response (or delay onset if there was no response), (3) short-delay feedback (with rewarded trials modeled as 1 and nonrewarded trials as −1), and (4) long-delay feedback (with rewarded trials modeled as 1 and nonrewarded as −1). This design meant that main effects of the feedback regressors would highlight regions of the brain that preferentially respond to rewarding versus nonrewarding feedback. Thus, combinations and contrasts of these two regressors can be used to highlight regions of the brain that are sensitive to reward across delays and also regions that are delay-sensitive.

We used a model-based approach to locate neural correlates of RPEs and surprise (|RPE|) in the brain. To do this, we fit RL models to participant choice data (described above) to obtain participant-specific learning and choice parameters. Then, we used these fitted parameters to simulate trial-by-trial RPEs based on the series of trials that each participant experienced. We included separate parametric regressors for RPEs on short-delay trials versus long-delay trials. In the main version of this model (presented in the main text), we included only rewarded trials in the analysis. Crucially, this allows us to measure true RPE responses that are not confounded with valence. In this model, we additionally included regressors for (1) stimulus appearance, (2) participant response, and (3) the appearance of the feedback. Thus, positive correlations with the RPE regressors should not be due to the appearance of feedback alone or generalized valence responses but rather with the trial-by-trial variation in

inferred RPE. In another version of this model (presented in Supplemental Fig. 3), we included both rewarded and nonrewarded trials. Here, we included an additional regressor for valence to ensure that correlations with the RPE regressors are not driven by general reward versus nonreward effects (though we note that the valence and RPE regressors are highly colinear, making this model suboptimal). We took a similar approach to examine correlates of surprise. In this case, the parametric regressor was the absolute value of RPEs. Finally, we also leveraged this computational framework to conduct an exploratory analysis examining reward anticipation signals in the cerebellum. In this analysis, we used a parametric regressor that had the Q value of the chosen stimulus for each trial, time-locked to when the stimuli were initially presented during the trial. All parametric regressors were z-scored before running the models.

We also conducted a whole-brain GLM with task performance as a covariate in the model to assess if RPE signals were related to behavior. Our primary goal was to assess whether cerebellar RPE activity was related to learning for the short-delay pairs. To that end, we quantified performance for the short-delay pairs by calculating the probability of selecting the higher reward probability shape across all presentations of these pairs. We included this metric as a covariate in the group-level analysis for temporally sensitive RPE signals (i.e., the contrast of RPE activity on rewarded trials on short- versus long-delay trials). We repeated this analysis with performance for long-delay pairs and again for the difference in performance between short and long-delay pairs.

In addition to our regressors of interest, we included motion and noise regressors in all GLMs. We extracted motion regressors from the fMRIprep pipeline and included six rigid body regressors, a DVARS regressor to account for overall motion, and regressors to scrub high motion (>0.5 mm) TRs. We included the first 10 components of the aCompCor regressors from fMRIprep to account for noise from white matter or cerebro-spinal fluid.

Runs were combined within subject before group-level analyses (FLAME 1). We used a cluster-forming threshold of $p < 0.001$ and a familywise error cluster-corrected threshold of $p < 0.001$, unless otherwise specified in the text. We additionally present results thresholded at 0 in Supplemental Fig. 4. For cerebellar-specific analyses, we cluster-corrected within the cerebellum. We used nilearn (version 0.8.1) to visualize whole-brain results and the SUIT toolbox (Diedrichsen, 2006; Diedrichsen et al., 2009; Diedrichsen and Zotow, 2015) to visualize cerebellar results on flat maps.

ROI analyses. We conducted additional analyses in a priori anatomical regions of interest (ROIs) to examine correlates of reward processing and RPE. We used anatomical ROIs from the Harvard–Oxford atlas to create ROI masks. We then extracted subjectwise average β-parameters from the second-level GLM analysis within each mask for each contrast of interest. We used nonparametric bootstrap hypothesis tests to test whether the measured effect was significantly different from zero in each ROI. We opted for these nonparametric tests as they are robust to statistical assumptions necessary for standard parametric tests. For each bootstrap iteration, we randomly sampled from the measured values with replacement to create a distribution. We repeated this process 1,000 times and computed p values and 95% confidence intervals by comparing the distribution means to the null value (zero in this case).

Psychophysiological interaction (PPI) analysis. We used PPI analyses (Gitelman et al., 2003) to examine task-dependent changes in functional correlations between the cerebellum and cerebrum during feedback processing. This approach examines whether functional coupling between a seed region and other regions of the brain is modulated by experimental factors. We first extracted a time series of activation from our two a priori cerebellar lobular ROIs (Crus I and Crus II) as the physiological regressors in this analysis. Then, we constructed the psychological regressor that identified feedback epochs in the task versus other trial epochs (e.g., stimulus appearance, delay, etc.) which was convolved with a canonical hemodynamic response function. Finally, we created a regressor that reflected the interaction of the physiological and psychological regressors. This regressor reveals regions of the brain that significantly covary with the seed region specifically during the epochs of interest.

Data and code availability statement

Data and code are available at https://github.com/jetrach/CBPE_RL_public.

Results

Human Crus I and Crus II preferentially respond to reward feedback

Participants ($N = 32$) performed a probabilistic RL task while undergoing fMRI. Participants observed two stimuli on each trial and used a button box to select the stimulus (left or right) that they thought was most likely to yield a reward (Fig. 1*a*). Their goal was to win points to earn a monetary bonus at the end of the task. After they made their choice, there was a delay period (short delay, 0.8 s; long delay, 3 s) before they received probabilistic reward feedback (nonrewarded trials, +0; rewarded trials, +1) which they could use to guide future choices. Participants saw four distinct pairs of stimuli within each run (see Materials and Methods and Supplemental Fig. 1). In each pair, one stimulus was associated with a 0.75 probability of reward and the other with a 0.25 probability of reward. Importantly, the location of the stimuli was randomized across trials, such that rewards were associated with the abstract stimulus choice rather than a specific motor action or spatial location. Our a priori ROIs were Crus I and Crus II in the cerebellum [following work in nonhuman primates (Sendhilnathan et al., 2020) and rodents (Heffley and Hull, 2019)] and canonical subcortical RL regions in the nucleus accumbens (NAc), caudate nucleus (Cd), and hippocampus (Hpc; Fig. 1*b*).

Participants performed the RL task well, showing significant learning across trial blocks [GLM, choose correct ~ iteration + (1|subject); iteration, $b = 0.026$; SE = 0.0037; $z = 6.93$; $p < 0.001$; Fig. 1*c*; see also Supplemental Fig. 2]. These learning data were fit with an RL model which was used for later model-based fMRI analyses (see Materials and Methods). Importantly, the plateau in the performance curve does not mean that participants were only actively learning early in the block—their performance plateaued at choosing the more rewarding stimulus on 73.8% of trials in the last five iterations of each stimulus pair rather than 100% which would be the optimal behavior. Additionally, because reward feedback was probabilistic, participants will continue to experience prediction errors throughout the block of trials.

At the group level, feedback delay did not significantly affect choice behavior or reaction times [p(choose higher value stimulus), $t_{(31)} = 0.51$; 95% CI = [−0.02, 0.04]; $p = 0.616$; RT, $t_{(31)} = 0.03$; 95% CI = [−0.01, 0.01]; $p = 0.976$; Fig. 1*d*]. Crucially, there was comparable gaze behavior (Fig. 1*e*; at feedback, N saccades, $t_{(28)} = 0.76$; 95% CI = [−0.05, 0.11]; $p = 0.451$) and physiological measures (Fig. 1*f*; rewarded vs nonrewarded trials, respiration, $t_{(23)} = -0.95$; 95% CI = [−0.37, 0.14]; $p = 0.353$; pulse, $t_{(18)} = 1.27$; 95% CI = [−0.25, 0.83]; $p = 0.274$) on rewarded versus nonrewarded trials, meaning that our key neural analyses would not be significantly confounded by these factors. (This is an important control, as confounding factors like eye movements can influence cerebellar activity.)

For our neural analyses, we first used GLMs to examine activity between rewarded and nonrewarded trials. In addition to canonical reward processing regions (Fig. 2*a*, **left panel**), we also observed significant activity in the lateral cerebellum around the Crus I/Crus II boundary that reflected greater feedback responses on rewarded versus nonrewarded trials (Fig. 2*a*,

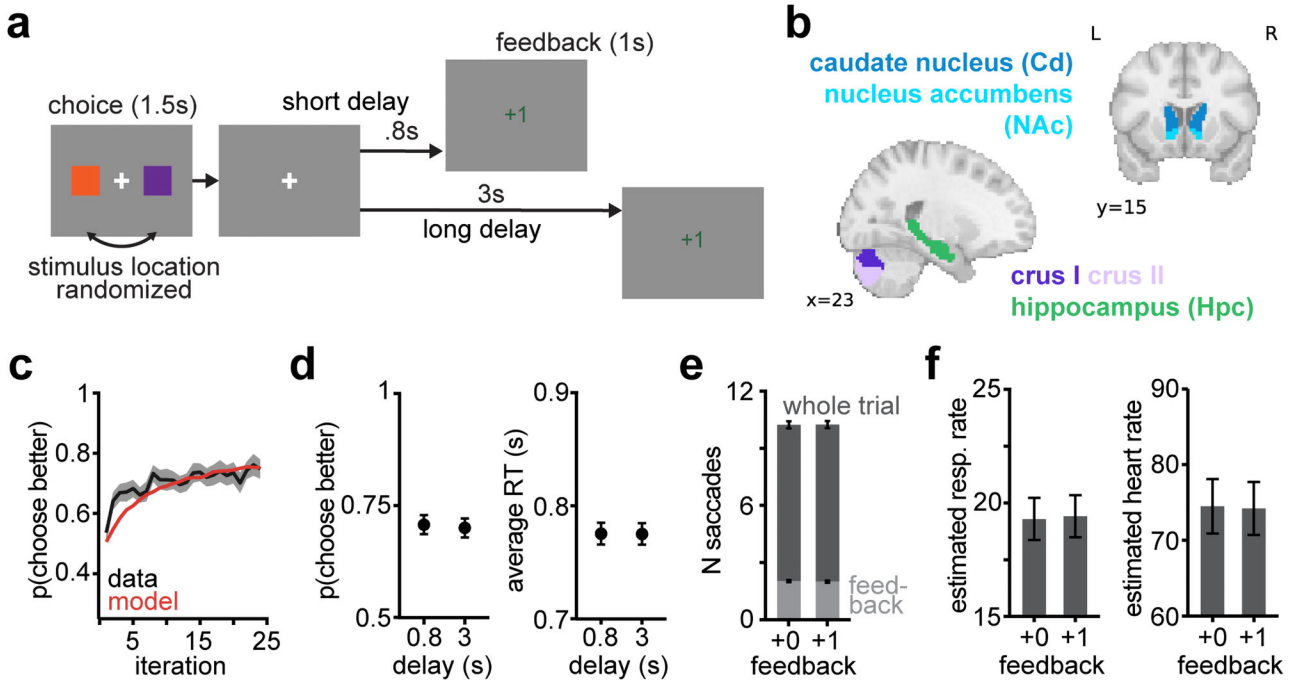

Figure 1. ***a***, Task schematic. Example stimuli pictured in Supplemental Fig. 1. ***b***, Anatomical ROI masks. ***c***, Learning curve and model fit. ***d***, Choice and RT performance on short- and long-delay trials. ***e***, Gaze on nonrewarded (+0) versus rewarded (+1) trials. Larger values depict data from whole trial and smaller values depict data during the feedback phase only. ***f***, Respiration and pulse rates on rewarded versus nonrewarded trials.

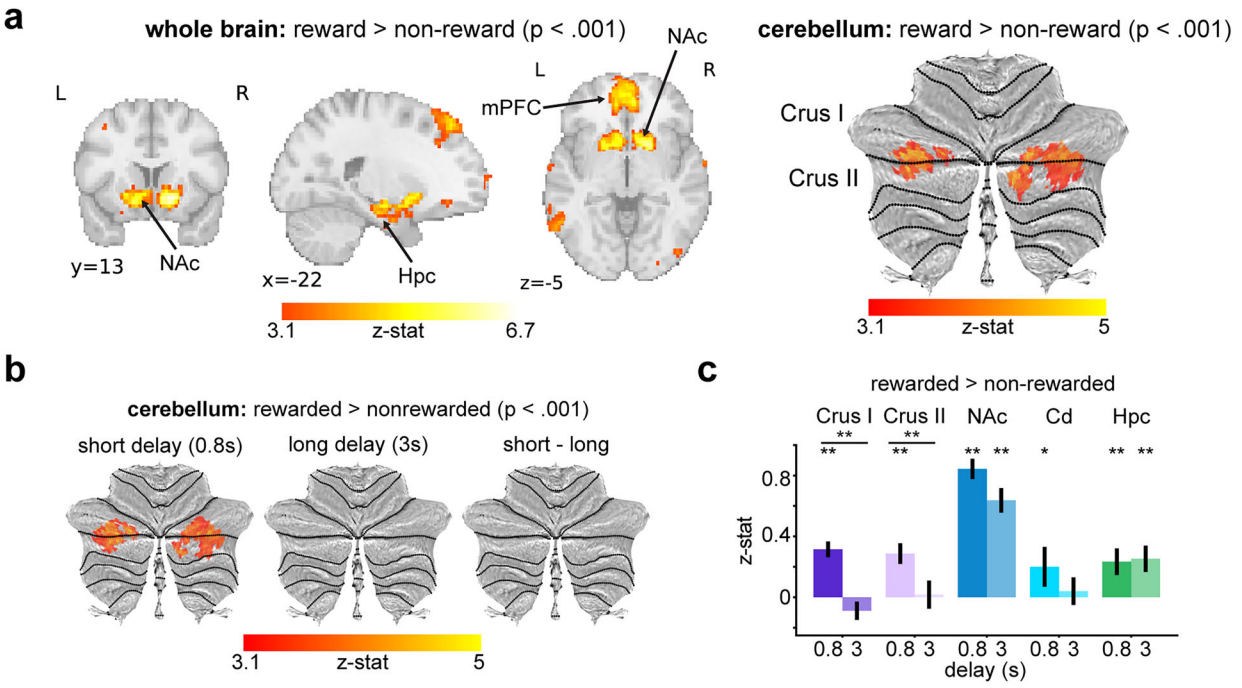

Figure 2. ***a***, Contrast of activity on rewarded versus nonrewarded trials at the point of feedback in the whole brain (left panel) and in cerebellum (right panel; plotted on cerebellar flat map; see Materials and Methods). Whole-brain result is cluster-corrected at $p < 0.001$. Cerebellar results are cluster-corrected in the cerebellum, $p < 0.001$. ***b***, Rewarded > nonrewarded contrast at short and long delays in the cerebellum (cluster-corrected in the cerebellum, $p < 0.001$). ***c***, Contrast of activity for rewarded > nonrewarded trials in cerebellar (Crus I and II), striatal (nucleus accumbens, NAc; caudate nucleus, Cd), and hippocampal (Hpc) ROIs. $^{**}p < 0.001$; $^{*}p < 0.05$.

right panel). The localization of these signals broadly corresponds with previously reported effects in both rodents (Heffley and Hull, 2019) and nonhuman primates (Sendhilnathan et al., 2020).

Cerebellar reward signals are temporally constrained

We next tested whether these cerebellar signals were affected by feedback delay. We hypothesized that the cerebellum would be specifically responsive to feedback on short-delay (0.8 s) versus

long-delay (3 s) trials, echoing temporal constraints on cerebellar processing in motor learning tasks (Kitazawa et al., 1995; Cheng et al., 2008; Brudner et al., 2016). We thus examined responses to rewards at short- and long-delay intervals separately. In line with our hypothesis, we found significant cerebellar responses to rewards on short-delay, but not long-delay, trials (Fig. 2*b*; Supplemental Fig. 5). Activity was primarily localized to Crus I and Crus II. The null findings on long-delay trials held even with a relaxed statistical threshold (cluster-corrected in the cerebellum, $p < 0.05$), indicating that the lack of significant reward responses at long delays was not simply a result of conservative thresholding.

We additionally asked whether this temporal sensitivity was specific to the cerebellum or was also reflected in other reward processing regions (NAc, Cd, and Hpc; Foerde and Shohamy, 2011). We extracted β values for the rewarded > nonrewarded contrast (Fig. 2*c*; see also Supplemental Fig. 5 for the main effects of reward vs nonreward feedback) from these ROIs at short and long delays. Corroborating the whole cerebellum results, the a priori anatomical cerebellar ROIs only exhibited significant positive responses at short delays (nonparametric bootstrap test, short delays, Crus I, $M = 0.32$; 95% CI = [0.19, 0.45]; $p < 0.001$; Crus II, $M = 0.29$; 95% CI = [0.16, 0.42]; $p < 0.001$; long delays, Crus I, $M = -0.09$; 95% CI = [−0.26, 0.10]; $p = 0.346$; Crus II, $M = 0.02$; 95% CI = [−0.15, 0.19]; $p = 0.844$). This pattern of results in the cerebellum was driven by significant positive activity during feedback on short-delay trials (Supplemental Fig. 5; nonparametric bootstrap test, Crus I, short-delay reward effect, $M = 0.51$; 95% CI = [0.33, 0.68]; $p < 0.001$; short-delay no–reward effect, $M = 0.28$; 95% CI = [0.12, 0.44]; $p < 0.001$; Crus II, short-delay reward effect, $M = 0.31$; 95% CI = [0.08, 0.53]; $p = 0.006$; short-delay no–reward effect, $M = 0.07$; 95% CI = [−0.10, 0.24]; $p = 0.045$) and negative activity on long-delay trials whose magnitude was invariant to the valence of the feedback (nonparametric bootstrap test, Crus I, long-delay reward effect, $M = -0.44$; 95% CI = [−0.64,−0.24]; $p < 0.001$; long-delay no–reward effect, $M = -0.43$; 95% CI = [−0.59,−0.25]; $p < 0.001$; Crus II, long-delay reward effect, $M = -0.11$; 95% CI = [−0.25,−0.05]; $p = 0.176$; long-delay no–reward effect, $M = -0.16$; 95% CI = [−0.31, −0.03]; $p = 0.010$).

In the striatum, caudate nucleus (Cd) showed a similar effect as the cerebellum (nonparametric bootstrap test, short delays, $M = 0.20$; 95% CI = [0.03, 0.36]; $p = 0.020$; long delays, $M = 0.04$; 95% CI = [−0.11, 0.20]; $p = 0.648$). However, the ventral striatum ROI (NAc) and hippocampus (Hpc) had significant responses at both short and long feedback delays (nonparametric bootstrap test, short delays, NAc, $M = 0.84$; 95% CI = [0.60, 1.08]; $p < 0.001$; Hpc, $M = 0.23$; 95% CI = [0.14, 0.33]; $p < 0.001$; long delays, NAc, $M = 0.64$; 95% CI = [0.45, 0.80]; $p < 0.001$; Hpc, $M = 0.25$; 95% CI = [0.13, 0.36]; $p < 0.001$). Only the cerebellar ROIs exhibited significantly stronger responses at short versus long feedback delays (short > long, Crus I, $M = 0.30$; 95% CI = [0.18, 0.44]; $p < 0.001$; Crus II, $M = 0.20$; 95% CI = [0.07, 0.33]; $p = 0.002$; Cd, $M = 0.12$; 95% CI = [−0.01, 0.25]; $p = 0.072$; NAc, $M = 0.15$; 95% CI = [−0.01, 0.31]; $p = 0.072$; Hpc, $M = -0.01$; 95% CI = [−0.11, 0.08]; $p = 0.826$). Furthermore, the difference in response to short- versus long-delay feedback was significantly larger in Crus I than in Cd or Hpc (Crus I, $M = 0.40$; NAc, $M = 0.21$; Cd, $M = 0.16$; Hpc, $M = -0.02$; Crus I vs NAc, 95% CI = [−0.07, 0.44]; $p = 0.142$; Crus I vs Cd, 95% CI = [0.07, 0.41]; $p = 0.006$; Crus I vs Hpc, 95% CI = [0.15,0.68]; $p = 0.002$) and in Crus II ($M = 0.27$) than in Hpc (Crus II vs NAc, 95% CI = [−0.2,0.31]; $p = 0.614$; Crus II vs Cd, 95% CI = [−0.1, 0.31]; $p = 0.298$; Crus II vs Hpc, 95% CI = [0.05,0.53]; $p = 0.016$). Taken together, these results provide evidence for temporally constrained reward processing in the human cerebellum during RL and also show that the delayed feedback did not globally abolish reward signals.

Model-based analyses reveal time-sensitive RPE computations in the human cerebellum

We next tested whether the human cerebellum tracked the core teaching signal in RL: RPE. We leveraged computational modeling to examine RPEs in the cerebellum (Fig. 3*a*). To do this, we fit RL models to each participant's behavior to obtain learning and choice parameters for each subject (see Materials and Methods; Fig. 1*c*; Supplemental Fig. 2). We then used these subject-specific parameters to simulate trial-by-trial RPEs based on a participant's unique sequence of trials and choices. The time course of these computationally inferred RPEs serves as a prediction about activity in regions of the brain that may encode RPE signals. We restricted our main analyses to RPEs on rewarded trials to avoid the strong collinearity with valence signals (see Materials and Methods; see Supplemental Fig. 3 for RPE results across rewarded and nonrewarded trials).

We found robust RPE responses at the boundary of Crus I and Crus II in the cerebellum, in addition to canonical RPE signals in the ventral striatum (Fig. 3*b*). Like reward-related activation, RPE responses were temporally constrained and only detectable on short-delay trials (Fig. 3*c*; Supplemental Fig. 3). Regions within Crus I and II, as well as caudate, showed significant positive RPE responses at short, but not long, delays (Fig. 3*d*; nonparametric bootstrap test, short delays, Crus I, $M = 0.65$; 95% CI = [0.47, 0.81]; $p < 0.001$; Crus II, $M = 0.45$; 95% CI = [0.28, 0.63]; $p < 0.001$; Cd, $M = 0.36$; 95% CI = [0.20, 0.52]; $p < 0.001$; long delays, Crus I, $M = -0.30$; 95% CI = [−0.46, −0.16]; $p < 0.001$; Crus II, $M = 0.01$; 95% CI = [−0.10, 0.12]; $p = 0.886$; Cd, $M = -0.10$; 95% CI = [−0.27, 0.07]; $p = 0.268$), while NAc and Hpc displayed robust RPE responses at both short and long delays (short delays, NAc, $M = 0.84$; 95% CI = [0.60, 1.08]; $p < 0.001$; Hpc, $M = 0.13$; 95% CI = [0.04, 0.23]; $p = 0.002$; long delays, NAc, $M = 0.54$; 95% CI = [0.39, 0.70]; $p < 0.001$; Hpc, $M = 0.22$; 95% CI = [0.11, 0.33]; $p < 0.001$; although stronger at short delays in NAc, short > long, $M = 0.44$; 95% CI = [0.23, 0.65]; $p < 0.001$; Foerde and Shohamy, 2011). The difference in RPE encoding for short- versus long-delay feedback was significantly larger in Crus I than all other ROIs (Crus I, $M = 0.95$; Crus II, $M = 0.44$; NAc, $M = 0.3$; Cd, $M = 0.45$; Hpc, $M = -0.08$; Crus I vs NAc, 95% CI = [0.4, 0.92]; $p < 0.001$; Crus I vs Cd, 95% CI = [0.26, 0.75]; $p < 0.001$; Crus I vs Hpc, 95% CI = [0.78, 1.28]; $p < 0.001$) and for Crus II versus Hpc (Crus II vs NAc, 95% CI = [−0.13,0.43]; $p = 0.316$; Crus II vs Cd, 95% CI = [−0.28, 0.26]; $p = 0.99$; Crus II vs Hpc, 95% CI = [0.29, 0.77]; $p < 0.001$). Thus, we found that the human cerebellum encodes prediction errors that are used for learning in a nonmotor learning task and that these teaching signals show similar temporal constraints as those observed in sensorimotor learning.

RPEs can be signed (i.e., sensitive to valence) or unsigned (i.e., only sensitive to outcome surprise, but not valence). Thus, we used the same approach as above to examine neural correlates of surprise (|RPE|). This analysis, however, did not reveal any significant clusters of activity in the cerebellum, even at a relaxed statistical threshold, suggesting that human cerebellar RPE activity is dominated by a reward-sensitive signal (see Discussion).

In addition, this model-based analysis framework can be used to examine reward anticipation signals in the brain. Reward anticipation signals can be operationalized as the trial-by-trial

Q value associated with the stimulus that the participant chose. Again, in this exploratory analysis, we identified regions of the brain that have activity that is positively correlated with a computationally inferred value, in this case trial-by-trial Q values. We did not observe any significant positive activity in the cerebellum related to stimulus Q values, although there was substantial activity in the canonical medial frontal regions that did correlate with stimulus Q values (Supplemental Fig. 3; Bartra et al., 2013). We return to this point in Discussion.

Linking cerebellar RPE activity to behavior

An additional question we asked is whether the apparent cerebellar involvement in RL at short (but not long) feedback delays covaries with behavioral performance in the task. We note that we did not see general performance differences between delay conditions at the group level: On average, participants showed comparable learning and RTs for stimulus pairs associated with short versus long feedback delays [p(choose higher value), $t_{(31)} = 0.51$; 95% CI = [−0.02, 0.04]; $p = 0.616$; RT, $t_{(31)} = 0.03$; 95% CI = [−0.01, 0.01]; $p = 0.976$; Fig. 1*d*]. However, we did observe individual differences in learning efficacy.

To investigate brain–behavior correlations, we compared RPE signals for short- versus long-delay trials in a whole-brain GLM and included RL performance in the short-delay condition as a participant-wise covariate in the model. In other words, this analysis tests for voxels in the brain where stronger temporally sensitive RPE responses covary with how well a participant performed in the short-delay condition. Strikingly, this whole-brain analysis revealed significant brain–behavior correlations only within the cerebellum: The analysis revealed a single significant cluster of activity on the Crus I/II boundary (albeit at a slightly relaxed statistical threshold for the whole-brain case; Fig. 4; see also Supplemental Fig. 4). As a control, we also repeated this analysis using performance on the long-delay condition and the difference between performance on the short- versus long-delay conditions as covariates, and observed no significant clusters. To visualize this result more clearly, we used a mask of voxels that showed this pattern in the cerebellum and plotted activity in this region for each subject against their performance at the

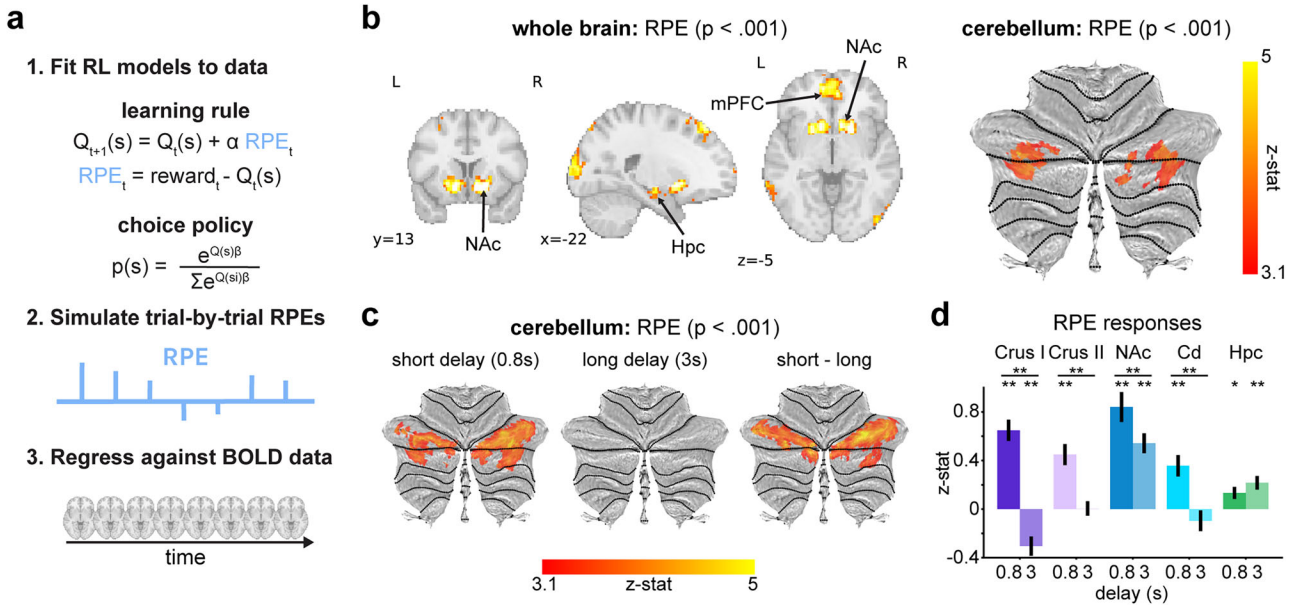

Figure 3. ***a***, Schematic of model-based fMRI approach. ***b***, RPE responses (whole brain cluster-corrected, $p < 0.001$) and cerebellum (cluster-corrected in the cerebellum, $p < 0.001$). ***c***, RPE response on rewarded trials at short and long feedback delays (cluster-corrected in the cerebellum, $p < 0.001$). ***d***, RPE responses at short and long delays in cerebellar and other subcortical ROIs. $^{**}p < 0.001$; $^{*}p < 0.05$.

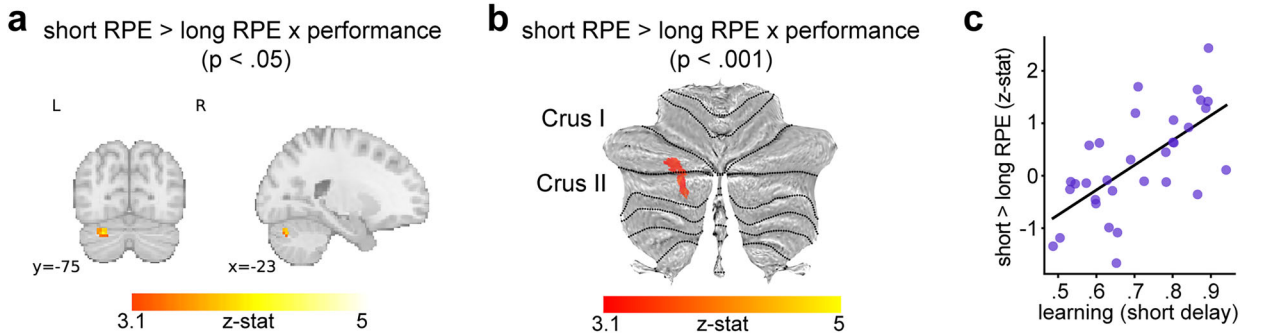

Figure 4. ***a***, Whole-brain analysis showing voxels where short RPE > long RPE contrast covaries with performance on short-delay trials (cluster-corrected in the whole brain, $p < 0.05$). ***b***, Same analysis restricted to the cerebellum (cluster-corrected in the cerebellum, $p < 0.001$). ***c***, Visualization of the correlation between activity in the significant cluster in the cerebellum from panel ***b*** and learning performance for short-delay pairs.

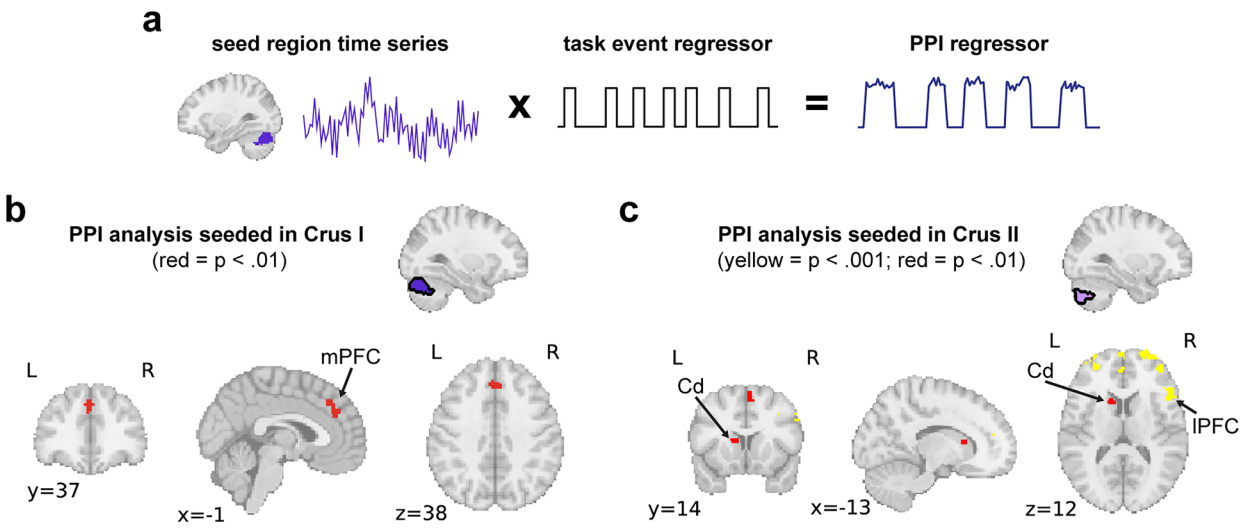

Figure 5. ***a***, Illustration of PPI analysis. ***b***, PPI analysis results seeded in Crus I (red regions cluster-corrected at $p < 0.01$). ***c***, PPI analysis results seeded in Crus II (red cluster-corrected at $p < 0.01$; yellow at $p < 0.001$).

short-delay condition (Fig. 4*c*). Although this analysis is correlational, it suggests that better RL (in the short-delay condition) was related to stronger temporally sensitive cerebellar RPEs, thus linking our cerebellar results to behavioral outcomes.

Increased connectivity between cerebellar ROIs and cerebrum during feedback

The cerebellum has bidirectional connections to much of the cerebral RL circuitry, including the basal ganglia and prefrontal cortex (Bostan et al., 2013; Bostan and Strick, 2018). Do cerebellar responses to RL feedback reflect connectivity with a wider cerebellar–striatal–frontal functional circuit? To address this question, we asked which regions of the cerebrum might be functionally coupled with the cerebellum during RL, specifically during feedback processing. We conducted exploratory PPI analyses (Di et al., 2021; Fig. 5*a*) seeded in our two a priori anatomical cerebellar ROIs (Crus I and Crus II) to measure if and how functional correlations between the cerebellum and the rest of the brain increased during RL feedback above and beyond other task phases.

We observed significant activity in the medial prefrontal cortex reflecting robust feedback-sensitive functional connectivity with cognitive regions of the cerebellum during RL, albeit at a slightly relaxed statistical threshold in the case of Crus I (Fig. 5*b*,*c*). Crus II also showed RL feedback-related functional correlations with lateral PFC regions. Furthermore, at the relaxed statistical threshold, we saw significant connectivity between Crus II and the caudate nucleus (Fig. 5*c*), the subcortical region that showed similar temporal sensitivity for RPE responses. These results suggest that nonmotor reward feedback increases functional coupling between the cerebellum and core corticostriatal circuits that support RL.

Discussion

Here, we present evidence of temporally constrained cerebellar involvement in RL in humans. We found that cognitive regions of the human cerebellum (clusters spanning Crus I and II) responded to rewards and encoded RPEs, extending recent work in model organisms (Wagner et al., 2017; Heffley and Hull, 2019; Kostadinov et al., 2019; Larry et al., 2019; Sendhilnathan et al., 2020). These cerebellar RL responses were sensitive to feedback timing, as in motor learning (Smith, 1968; Ivry and Keele, 1989; Gerwig et al., 2008). Furthermore, the degree of time-sensitive RPE encoding covaried with performance on short-delay trials, connecting cerebellar RPEs to the efficacy of RL. Finally, functional connectivity analyses revealed communication between the cerebellum and corticostriatal circuitry during feedback (Hoshi et al., 2005; Bostan et al., 2010; Bostan and Strick, 2018; Carta et al., 2019). These results implicate the cerebellum in prediction error-based learning beyond supervised sensorimotor learning and highlight the importance of examining cerebellar responses in cognitive neuroscience (Wang et al., 2025).

We note that our results are not the first to report prediction error signals in the human cerebellum (O'Doherty et al., 2003; Lee et al., 2025) or to suggest that the basal ganglia and cerebellum work collaboratively (Bostan and Strick, 2010). Previous work, however, did not necessarily isolate the signals we see here: First, in previous work, rewards were typically attached to specific motor actions, making it unclear whether cerebellar responses were driven by RL versus sensorimotor learning (Wagner et al., 2017; Sendhilnathan et al., 2024). Similarly, in other studies, reward or punishment feedback had significant sensorimotor components, including shocking of the skin and the ingestion of juice (Maschke et al., 2002; O'Doherty et al., 2003). Such feedback is likely to drive sensorimotor learning processes in the cerebellum and complicates interpretations of cerebellar prediction error responses.

The fact that cerebellar RL signals were temporally sensitive could also explain why similar signals have not been reported consistently in the literature. fMRI studies are typically designed with long delays between stimuli to accommodate the slow hemodynamic response or suprasecond TR measurements, which could blunt cerebellar learning signals. Additionally, much human fMRI work does not report what exact regions of the cerebellum are involved. Functional mapping of the cerebellar cortex has revealed a mosaic of functional regions, mirroring the complexity of the cerebral cortex (King et al., 2019; Nettekoven et al., 2024; Saadon-Grosman et al., 2024). These advances make specific localization of cerebellar signals increasingly important and de-emphasize the functional importance of

lobular boundaries. Notably, precision mapping research has characterized multiple functional subregions that are interconnected with distinct cortical and striatal networks (Habas et al., 2009; Nettekoven et al., 2024; Saadon-Grosman et al., 2024). The regions of the cerebellum where we see robust RL signals correspond to social–linguistic–spatial and demand regions (Guell et al., 2018; King et al., 2019; Nettekoven et al., 2024). These regions also exhibit resting state connectivity with default mode areas and wider frontoparietal and corticostriatal networks (Habas et al., 2009; Dobromyslin et al., 2012; Guell et al., 2018; Boonstra, 2025).

Importantly, we were able to draw from the animal literature to make broad a priori predictions about localization. We found correspondence between RL-sensitive regions reported in research with model organisms and our human participants, despite key differences in experimental design. While model organisms afford exceptional access to recording input and output signals in the cerebellum, one advantage of testing humans is that they can engage in rapid RL over a matter of minutes, respond to verbal instructions, and tolerate higher task variability across trials (e.g., multiple feedback delays). We think this work highlights the preservation of cerebellar function across model organisms and emphasizes the importance of cross talk between human and animal researchers. This idea is particularly relevant as cerebellar research becomes focused on higher-order cognitive functions (Strick et al., 2009; Buckner, 2013; Manto et al., 2024) and as cognitive neuroscientists increasingly consider cerebellar nonmotor functions.

One highlight of this work is that we observed consistent evidence of temporal constraints on cerebellar RL feedback processing; cerebellar encoding of reward and RPE was essentially abolished when feedback was delayed. The idea that the cerebellum is primarily involved in subsecond coordination, associative learning, and timing is well established in the motor domain (Smith, 1968; Ivry and Keele, 1989; Kitazawa et al., 1995; Gerwig et al., 2008; Ohmae and Medina, 2015; Barri et al., 2022). This preference makes sense for a brain region originally adapted for short-timescale sensorimotor control and prediction (Cisek, 2022). At a mechanistic level, properties of granule cell firing and short-term plasticity at mossy fiber-granule cell synapses have been proposed to support subsecond timing computations characteristic of the cerebellum (Yamazaki and Tanaka, 2009; Kennedy et al., 2014; Narain et al., 2018; Barri et al., 2022). These mechanisms can induce timing-specific activity patterns that are then used by Purkinje cells (the main output neurons of the cerebellum) to refine actions or cognitive computations. There are likely biological limits that constrain what time intervals can be precisely learned by these mechanisms (Raymond and Medina, 2018), providing a mechanistic account for why the cerebellum is attuned to short-timescale feedback.

Our model-based analysis approach also allowed for exploratory analyses examining reward anticipation signals. Recent work in animal models has found reward anticipation signals during RL in analogous regions (Wagner et al., 2017; Heffley and Hull, 2019; Kostadinov et al., 2019; Larry et al., 2019). For example, Larry et al. (2019) observed climbing fiber responses that scaled with expected reward during the cue phase of an RL task (i.e., classic temporal–difference learning; Sutton and Barto, 1998), echoing observations in the domain of eyeblink conditioning (Ohmae et al., 2017). Similarly, Wagner et al. (2017) reported granule cell activity ramping toward an expected reward. In contrast, we did not see significant positive cerebellar activity correlating with trialwise *Q* values (Supplemental Fig. 3). Why the discrepancy between our results and recent work in animal models?

The most likely explanation is the major differences in methodological approaches between human fMRI study and electrophysiological approaches. BOLD signals in the cerebellum primarily track activity in the mossy fiber-granule cell input pathway to the cerebellum (Shahshahani et al., 2024). Thus, climbing fiber signals at choice may be a relatively weak contribution to the BOLD signal versus combined climbing fiber and mossy fiber-granule cell activity during feedback. As for ramping signals (Wagner et al., 2017), these may be hard to separate from feedback-related responses given the low temporal resolution of fMRI. Finally, while model-based analyses are extremely useful for making predictions about RPE activity, it is important to remember that these RPE time courses are inferred from a necessarily simplified model of RL and model fit may vary considerably across participants.

Our connectivity analyses implicate communication between cognitive regions of the cerebellum and neocortical and subcortical regions during RL feedback processing. Some regions of medial PFC also emerged in our reward processing and RPE analyses, perhaps suggesting that cofluctuations between the cerebellum and these regions are driven by the receipt of rewards. Lateral PFC, in contrast, was not a reward-sensitive region in other analyses. Communication between cerebellum and lateral PFC could reflect general feedback processing, perhaps updating predictions based on new feedback. One caveat is that PPI cannot distinguish between correlations in activity due to functional connectivity versus high correlations in activity driven by alternative sources of shared input (Friston et al., 1997; Sanchez-Romero et al., 2019). The existence of anatomical connections between the cerebellum and prefrontal regions via the dentate nucleus and thalamus (Kelly and Strick, 2003) are consistent with a functional coupling explanation; however, future work with causal manipulations is necessary to resolve this issue.

Although we are primarily measuring cerebellar input, communication likely involves bidirectional connections between the cerebellum and frontal and subcortical RL regions (Hoshi et al., 2005; Bostan and Strick, 2010, 2018; Wagner and Luo, 2020; Saadon-Grosman et al., 2024; Boonstra, 2025). A recent fMRI paper demonstrated how cerebellar–striatal connectivity might support reward-based motor learning (Lee et al., 2025); in contrast to our work, they found that cerebellar activity primarily tracked surprise (i.e., unsigned prediction error magnitude). Indeed, we did not find significant activation in the cerebellum tracking surprise. This discrepancy is likely due to significant task differences, where cerebellar signals in Lee et al. (2025) were associated with successful versus unsuccessful outcomes of a motor task, not a nonmotor RL task as presented here.

Overall, our results provide evidence of temporally constrained, behaviorally relevant RL signals in the human cerebellum. RPE and reward-related signals were prominent in cognitive regions of the cerebellum, corresponding with recent work in animal models. Connectivity results demonstrated functional correlations between functional subregions spanning Crus I and II and regions of the medial and lateral frontal cortex and caudate nucleus during RL. This work extends beyond the well-studied role of the cerebellum in prediction error-based supervised motor learning (Marr, 1969; Albus, 1971; Wolpert et al., 1998; Raymond and Medina, 2018; Hull, 2020) and contributes to growing evidence that the human cerebellum may support prediction error-based learning across multiple domains.

References

Albus JS (1971) A theory of cerebellar function. Math Biosci 10:25–61.

Avants BB, Epstein CL, Grossman M, Gee JC (2008) Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 12:26–41.

Avraham G, Taylor JA, Breska A, Ivry RB, McDougle SD (2022) Contextual effects in sensorimotor adaptation adhere to associative learning rules. eLife 11:e75801.

Barri A, Wiechert MT, Jazayeri M, DiGregorio DA (2022) Synaptic basis of a sub-second representation of time in a neural circuit model. Nat Commun 13:7902.

Bartra O, McGuire JT, Kable JW (2013) The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage 76:412–427.

Behzadi Y, Restom K, Liau J, Liu TT (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37:90–101.

Billauer E (2009) Peakdet (Version 3.4.05) [MATLAB].

Boonstra JT (2025) The cerebellar connectome. Behav Brain Res 482:115457.

Boorman ED, Behrens TE, Rushworth MF (2011) Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. PLoS Biol 9:e1001093.

Bostan AC, Strick PL (2010) The cerebellum and basal ganglia are interconnected. Neuropsychol Rev 20:261–270.

Bostan AC, Strick PL (2018) The basal ganglia and the cerebellum: nodes in an integrated network. Nat Rev Neurosci 19:338–350.

Bostan AC, Dum RP, Strick PL (2010) The basal ganglia communicate with the cerebellum. Proc Natl Acad Sci U S A 107:8452–8456.

Bostan AC, Dum RP, Strick PL (2013) Cerebellar networks with the cerebral cortex and basal ganglia. Trends Cogn Sci (Regul Ed) 17:241–254.

Brudner SN, Kethidi N, Graeupner D, Ivry RB, Taylor JA (2016) Delayed feedback during sensorimotor learning selectively disrupts adaptation but not strategy use. J Neurophysiol 115:1499–1511.

Buckner RL (2013) The cerebellum and cognitive function: 25 years of insight from anatomy and neuroimaging. Neuron 80:807–815.

Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BTT (2011) The organization of the human cerebellum estimated by intrinsic functional connectivity. J Neurophysiol 106:2322–2345.

Buhusi CV, Meck WH (2005) What makes us tick? Functional and neural mechanisms of interval timing. Nat Rev Neurosci 6:755–765.

Butcher PA, Ivry RB, Kuo S-H, Rydz D, Krakauer JW, Taylor JA (2017) The cerebellum does more than sensory prediction error-based learning in sensorimotor adaptation tasks. J Neurophysiol 118:1622–1636.

Carta I, Chen CH, Schott AL, Dorizan S, Khodakhah K (2019) Cerebellar modulation of the reward circuitry and social behavior. Science 363:6424.

Cheng DT, Disterhoft JF, Power JM, Ellis DA, Desmond JE (2008) Neural substrates underlying human delay and trace eyeblink conditioning. Proc Natl Acad Sci U S A 105:8108–8113.

Chettih SN, McDougle SD, Ruffolo LI, Medina JF (2011) Adaptive timing of motor output in the mouse: the role of movement oscillations in eyelid conditioning. Front Integr Neurosci 5:1–11.

Cisek P (2022) Evolution of behavioural control from chordates to primates. Philos Trans R Soc Lond B Biol Sci 377:20200522.

Cohen JD, et al. (2017) Computational approaches to fMRI analysis. Nat Neurosci 20:304–313.

Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 9:179–194.

Di X, Zhang Z, Biswal BB (2021) Understanding psychophysiological interaction and its relations to beta series correlation. Brain Imaging Behav 15:958–973.

Diedrichsen J (2006) A spatially unbiased atlas template of the human cerebellum. Neuroimage 33:127–138.

Diedrichsen J, McDougle SD (2026) How does the cerebellum contribute to cognitive functions? PLoS Biol 24:e3003688.

Diedrichsen J, Zotow E (2015) Surface-based display of volume-averaged cerebellar imaging data. PLoS One 10:e0133402.

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, Verstynen T, Schlerf J, Wiestler T (2010). Advances in functional imaging of the human Cerebellum. Curr Opin Neurol 23:382–387.

Diedrichsen J, King M, Hernandez-Castillo C, Sereno M, Ivry RB (2019) Universal transform or multiple functionality? Understanding the contribution of the human cerebellum across task domains. Neuron 102:918–928.

Dobromyslin VI, Salat DH, Fortier CB, Leritz EC, Beckmann CF, Milberg WP, McGlinchey RE (2012) Distinct functional networks within the cerebellum and their relation to cortical systems assessed with independent component analysis. Neuroimage 60:2073–2085.

Drepper J, Timmann D, Kolb FP, Diener HC (1999) Non-motor associative learning in patients with isolated degenerative cerebellar disease. Brain 122:87–97.

Esteban O, Ross B, Markiewicz CJ, Berleant SL, Moodie C, Ma F, Isik AI (2018) fMRIPrep 23.2.1. [Computer software].

Esteban O, et al. (2019) fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat Methods 16:111–116.

Foerde K, Shohamy D (2011) The role of the basal ganglia in learning and memory: insight from Parkinson's disease. Neurobiol Learn Mem 96:624–636.

Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE (2007) Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc Natl Acad Sci U S A 104:16311–16316.

Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997) Psychophysiological and modulatory interactions in neuroimaging. Neuroimage 6:218–229.

Garrison J, Erdeniz B, Done J (2013) Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. Neurosci Biobehav Rev 37:1297–1310.

Gerwig M, Eßer AC, Guberina H, Frings M, Kolb FP, Forsting M, Aurich V, Beck A, Timmann D (2008) Trace eyeblink conditioning in patients with cerebellar degeneration: comparison of short and long trace intervals. Exp Brain Res 187:85–96.

Gitelman DR, Penny WD, Ashburner J, Friston KJ (2003) Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. Neuroimage 19:200–207.

Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, Ghosh SS (2011) Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. Front Neuroinform 5:1–15.

Gorgolewski KJ, Esteban O, Markiewwicz CJ, Ziegler E, Ellis DG, Notter MP, Jarecka D (2018) Nipype [Computer software].

Greve DN, Fischl B (2009) Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48:63–72.

Guell X, Schmahmann JD, Gabrieli JD, Ghosh SS (2018) Functional gradients of the cerebellum. eLife 7:e36652.

Habas C, Kamdar N, Nguyen D, Prater K, Beckmann CF, Menon V, Greicius MD (2009) Distinct cerebellar contributions to intrinsic connectivity networks. J Neurosci 29:8586–8594.

Heffley W, Hull C (2019) Classical conditioning drives learned reward prediction signals in climbing fibers across the lateral cerebellum. eLife 8:8–10.

Heffley W, Song EY, Xu Z, Taylor BN, Hughes MA, McKinney A, Joshua M, Hull C (2018) Coordinated cerebellar climbing fiber activity signals learned sensorimotor predictions. Nat Neurosci 21:1431–1441.

Hoshi E, Tremblay L, Féger J, Carras PL, Strick PL (2005) The cerebellum communicates with the basal ganglia. Nat Neurosci 8:1491–1493.

Hull C (2020) Prediction signals in the cerebellum: beyond supervised motor learning. eLife 9:1–22.

Huvermann DM, et al. (2025) The cerebellum contributes to prediction error coding in reinforcement learning in humans. J Neurosci 45:e1972242025.

Iadecola C, Yang G, Ebner TJ, Chen G (1997) Local and propagated vascular responses evoked by focal synaptic activity in cerebellar cortex. J Neurophysiol 78:651–659.

Ito M (2008) Control of mental activities by internal models in the cerebellum. Nat Rev Neurosci 9:304–313.

Ivry RB, Keele SW (1989) Timing functions of the cerebellum. J Cogn Neurosci 1:136–152.

Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17:825–841.

Kelly RM, Strick PL (2003) Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. J Neurosci 23:8432–8444.

Kennedy A, Wayne G, Kaifosh P, Alviña K, Abbott LF, Sawtell NB (2014) A temporal basis for predicting the sensory consequences of motor commands in an electric fish. Nat Neurosci 17:416–422.
King M, Hernandez-Castillo CR, Poldrack RA, Ivry RB, Diedrichsen J (2019) Functional boundaries in the human cerebellum revealed by a multi-domain task battery. Nat Neurosci 22:1371–1378.
King M, Shahshahani L, Ivry RB, Diedrichsen J (2023) A task-general connectivity model reveals variation in convergence of cortical inputs to functional regions of the cerebellum. eLife 12:e81511.
Kitazawa S, Kohno T, Uka T (1995) Effects of delayed visual information on the rate and amount of prism adaptation in the human. J Neurosci 15: 7644–7652.
Klein A, et al. (2017) Mindboggling morphometry of human brains. PLoS Comput Biol 13:e1005350.
Kostadinov D, Häusser M (2022) Reward signals in the cerebellum: origins, targets, and functional implications. Neuron 110:1290–1303.
Kostadinov D, Beau M, Pozo MB, Häusser M (2019) Predictive and reactive reward signals conveyed by climbing fiber inputs to cerebellar Purkinje cells. Nat Neurosci 22:950–962.
Kruithof ES, Klaus J, Schutter DJLG (2023) The human cerebellum in reward anticipation and outcome processing: an activation likelihood estimation meta-analysis. Neurosci Biobehav Rev 149:105171.
Larry N, Yarkoni M, Lixenberg A, Joshua M (2019) Cerebellar climbing fibers encode expected reward size. eLife 8:1–16.
Lee JL, Casamento-Moran A, Bastian AJ, Cullen KE, Chib VS (2025) Striatal and cerebellar interactions during reward-based motor performance. Proc Natl Acad Sci U S A 122:e2503373122.
Leiner HC, Leiner AL, Dow RS (1986) Does the cerebellum contribute to mental skills? Behav Neurosci 100:443–454.
Leong YC, Radulescu A, Daniel R, DeWoskin V, Niv Y (2017) Dynamic interaction between reinforcement learning and attention in multidimensional environments. Neuron 93:451–463.
Manto M, et al. (2024) Consensus paper: cerebellum and reward. Cerebellum 23:2169–2192.
Marr D (1969) A theory of cerebellar cortex. J Physiol 202:437–470.
Maschke M, Schugens M, Kindsvater K, Drepper J, Kolb FP, Diener H-C, Daum I, Timmann D (2002) Fear conditioned changes of heart rate in patients with medial cerebellar lesions. J Neurol Neurosurg Psychiatry 72:116–118.
McDougle SD, Boggess MJ, Crossley MJ, Parvin D, Ivry RB, Taylor JA (2016) Credit assignment in movement-dependent reinforcement learning. Proc Natl Acad Sci U S A 113:6797–6802.
McDougle SD, Tsay JS, Pitt B, King M, Saban W, Taylor JA, Ivry RB (2022) Continuous manipulation of mental representations is compromised in cerebellar degeneration. Brain 145:4246–4263.
Middleton FA, Strick PL (1994) Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. Science 266:458–461.
Narain D, Remington ED, Zeeuw CID, Jazayeri M (2018) A cerebellar mechanism for learning prior distributions of time intervals. Nat Commun 9: 469.
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 (2023) The role of the cerebellum in learning to predict reward: evidence from cerebellar ataxia. Cerebellum 23:1355–1368.
O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. Neuron 38:329–337.
Ohmae S, Medina JF (2015) Climbing fibers encode a temporal-difference prediction error during cerebellar learning in mice. Nat Neurosci 18: 1798–1803.
Ohmae S, Kunimatsu J, Tanaka M (2017) Cerebellar roles in self-timing for sub- and supra-second intervals. J Neurosci 37:3511–3522.
Patriat R, Reynolds RC, Birn RM (2017) An improved model of motion-related signal changes in fMRI. Neuroimage 144:74–82.
Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE (2014) Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 84:320–341.
Raymond JL, Medina JF (2018) Computational principles of supervised learning in the cerebellum. Annu Rev Neurosci 41:233–253.
Saadon-Grosman N, Du J, Kosakowski HL, Angeli PA, DiNicola LM, Eldaief MC, Buckner RL (2024) Within-individual organization of the human cognitive cerebellum: evidence for closely juxtaposed, functionally specialized regions. Sci Adv 10:eadq4037.
Sanchez-Romero R, Ramsey JD, Zhang K, Glymour MRK, Huang B, Glymour C (2019) Estimating feedforward and feedback effective connections from fMRI time series: assessments of statistical methods. Netw Neurosci 3: 274–306.
Satterthwaite TD, et al. (2013) An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. Neuroimage 64:240–256.
Schlerf J, Ivry RB, Diedrichsen J (2012) Encoding of sensory prediction errors in the human cerebellum. J Neurosci 32:4913–4922.
Schmahmann JD (2019) The cerebellum and cognition. Neurosci Lett 688: 62–75.
Schmahmann JD, Sherman JC (1998) The cerebellar cognitive affective syndrome. Brain 121:561–579.
Sendhilnathan N, Ipata AE, Goldberg ME (2020) Neural correlates of reinforcement learning in mid-lateral cerebellum. Neuron 106:188–198.e5.
Sendhilnathan N, Bostan AC, Strick PL, Goldberg ME (2024) A cerebro-cerebellar network for learning visuomotor associations. Nat Commun 15:2519.
Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, Frackowiak RS (2004) Temporal difference models describe higher-order learning in humans. Nature 429:664–667.
Shahshahani L, King M, Nettekoven C, Ivry RB, Diedrichsen J (2024) Selective recruitment of the cerebellum evidenced by task-dependent gating of inputs. eLife 13:RP96386.
Smith MC (1968) CS-US interval and US intensity in classical conditioning of the rabbit's nictitating membrane response. J Comp Physiol Psychol 66: 679–687.
Sokolov AA, Miall RC, Ivry RB (2017) The cerebellum: adaptive prediction for movement and cognition. Trends Cogn Sci (Regul Ed) 21:313–332.
Stoodley CJ (2016) The cerebellum and neurodevelopmental disorders. Cerebellum 15:34–37.
Stoodley CJ, Valera EM, Schmahmann JD (2012) Functional topography of the cerebellum for motor and cognitive tasks: an fMRI study. Neuroimage 59:1560–1570.
Strick PL, Dum RP, Fiez JA (2009) Cerebellum and nonmotor function. Annu Rev Neurosci 32:413–434.
Striemer CL, Chouinard PA, Goodale MA, de Ribaupierre S (2015) Overlapping neural circuits for visual attention and eye movements in the human cerebellum. Neuropsychologia 69:9–21.
Sutton RS, Barto AG (1998) *Introduction to reinforcement learning*. Cambridge, MA: The MIT Press.
Timmann D, Drepper J, Frings M, Maschke M, Richter S, Gerwig M, Kolb FP (2010) The human cerebellum contributes to motor, emotional and cognitive associative learning. A review. Cortex 46:845–857.
Trach JE, deBettencourt MT, Radulescu A, McDougle SD (2025) Rewards transiently and automatically enhance sustained attention. J Exp Psychol Gen 154:1063–1079.
Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC (2010) N4ITK: improved N3 bias correction. IEEE Trans Med Imaging 29:1310–1320.
Van Overwalle F, Baetens K, Mariën P, Vandekerckhove M (2014) Social cognition and the cerebellum: a meta-analysis of over 350 fMRI studies. Neuroimage 86:554–572.
Van Overwalle F, et al. (2020) Consensus paper: cerebellum and social cognition. Cerebellum 19:833–868.
Wagner MJ, Luo L (2020) Neocortex–cerebellum circuits for cognitive processing. Trends Neurosci 43:42–54.
Wagner MJ, Kim TH, Savall J, Schnitzer MJ, Luo L (2017) Cerebellar granule cells encode the expectation of reward. Nature 544:96–100.
Wang B, LeBel A, D'Mello AM (2025) Ignoring the cerebellum is hindering progress in neuroscience. Trends Cogn Sci (Regul Ed) 29:318–330.
Wolpert DM, Miall RC, Kawato M (1998) Internal models in the cerebellum. Trends Cogn Sci (Regul Ed) 2:338–347.
Yamazaki T, Tanaka S (2009) Computational models of timing mechanisms in the cerebellar granular layer. Cerebellum 8:423–432.
Zhang Y, Brady M, Smith S (2001) Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm. IEEE Trans Med Imaging 20:45–57.