

Mixed Selectivity in the Cerebellar Purkinje-Cell Response during Visuomotor Association Learning

 Naveen Sendhilnathan,^{1,2,3,5} Michael E. Goldberg,^{2,3,4,5,6*} and Anna E. Ipata^{2,3,5*}

¹Doctoral Program in Neurobiology and Behavior, Columbia University, New York 10027, New York, ²Department of Neuroscience, Columbia University, New York 10027, New York, ³Mahoney Center for Brain and Behavior Research, Columbia University, New York 10032, New York, ⁴Kavli Institute for Brain Science, Columbia University, New York 10027, New York, ⁵Zuckerman Mind, Brain, and Behavior Institute, Columbia University, New York 10027, New York, and ⁶Department of Neurology, Psychiatry, and Ophthalmology, Columbia University College of Physicians and Surgeons, New York 10032, New York

Although the cerebellum has been traditionally considered to be exclusively involved in motor control, recent anatomic and clinical studies show that it also has a role in reward-processing. However, the way in which the movement-related and the reward-related neural activity interact at the level of the cerebellar cortex and contribute toward learning is still unclear. Here, we studied the simple spike activity of Purkinje cells in the mid-lateral cerebellum when 2 male monkeys learned to associate a right or left-hand movement with one of two visual symbolic cues. These cells had distinctly different discharge patterns between an overtrained symbol–hand association and a novel symbol–hand association, responding in association with the movement of both hands, although the kinematics of the movement did not change between the two conditions. The activity change was not related to the pattern of the visual symbols, the movement kinematics, the monkeys' reaction times, or the novelty of the visual symbols. The simple spike activity changed throughout the learning process, but the concurrent complex spikes did not instruct that change. Although these neurons also have reward-related activity, the reward-related and movement-related signals were independent. We suggest that this mixed selectivity may facilitate the flexible learning of difficult reinforcement learning problems.

Key words: cerebellum; electrophysiology; primates; Purkinje cells; reinforcement learning; reward

Significance Statement

The cerebellum receives both motor-related and reward-related information. However, it is unclear how these two signals interact at the level of cerebellar cortex and contribute to learning nonmotor skills. Here we show that in the mid-lateral cerebellum, the reward information is encoded independently from the motor information such that during reward-based learning, only the reward information carried by the Purkinje cells inform learning while the motor information remains unchanged with learning.

Introduction

Historically, the cerebellum has been considered to be exclusively involved in the acquisition of motor skills (De Zeeuw and Ten Brinke, 2015), motor learning, and gain adaptation (Bastian and

Lisberger, 2021). However, recent converging evidence (Heffley et al., 2018; Heffley and Hull, 2019; Kostadinov et al., 2019; Larry et al., 2019; Sendhilnathan et al., 2020, 2021) strongly suggests that the cerebellum also processes reward-related information. In particular, the lateral area of Lobule VII of the cerebellar hemisphere, near Crus I and II, is a region that has extensive reciprocal connections with other reward processing areas of the brain, including PFC and basal ganglia (Tedesco et al., 2011; Koziol et al., 2014; Caligiore et al., 2017). When monkeys learn to associate an arbitrary symbol with a well-learned movement in a reinforcement learning paradigm, Purkinje cell (P-cell) simple spikes produce an error signal, which decreases as the monkeys learn the task (Sendhilnathan et al., 2020). Transient inactivation of this area impaired the monkeys' ability to learn a new set of associations (Sendhilnathan and Goldberg, 2020).

Although the cerebellum receives both motor-related (Giovannucci et al., 2017; Raymond and Medina, 2018) and reward-related information (Wagner et al., 2017), their

Received Aug. 31, 2021; revised Feb. 18, 2022; accepted Mar. 18, 2022.

Author contributions: N.S., M.E.G., and A.E.I. designed research; N.S. and A.E.I. performed research; N.S. contributed unpublished reagents/analytic tools; N.S. analyzed data; N.S. wrote the first draft of the paper; N.S. and M.E.G. edited the paper; N.S. and M.E.G. wrote the paper.

This work was supported by the Keck, Zegar Family; Dana Foundations; and the National Eye Institute R24EY-015634, R21 EY-020631, R01 EY-017039, R01-NS113078, and P30 EY-019007 to M.E.G. Raw data are available online at <https://doi.org/10.17632/22n9ps5rzv.1>. We thank Glen Duncan for highly creative and wonderful electronic assistance; John Caban and Matthew Hasday for superb machining; Dr. Girma Asfaw and Dr. Moshe Shalev for animal care; Whitney Thomas and Holly Cline for facilitating everything; and Dr. Andreea Boston for help with identifying the locations of our electrode tracks in the MRIs.

*M.E.G. and A.E.I. contributed equally to this work.

The authors declare no competing financial interests.

Correspondence should be addressed to Naveen Sendhilnathan at ns3046@columbia.edu.

<https://doi.org/10.1523/JNEUROSCI.1771-21.2022>

Copyright © 2022 the authors

interaction at the level of cerebellar cortex and their contribution to learning is still unclear. To study the nature of information encoded by the cerebellum during learning an arbitrary stimulus–response association, we compared the activity of P cells in Crus I and II while 2 monkeys were actively learning new visuomotor associations and while they performed an already well-learned, familiar visuomotor association task. We found that the simple spikes of hand-related P cells changed their activity profile significantly when we changed the visuomotor association from a well-learned to a novel (L) association, although the monkeys made the precisely same hand movement to report their choices. The pattern of the simple spike response continued to change throughout the learning process. This change in neural activity was independent of the kinematics of the movement, the hand used to report the choice, the complexity of the visual symbols, the reaction time of the animal, or the novelty of the visual symbols. Furthermore, this signal could be dissociated from the reward-related error signal, which described the result of the monkey's prior decision (Sendhilnathan et al., 2020). The concurrent complex spike activity was unlikely to have instructed this change in simple spike activity. Our results provide evidence that P cells in the lateral part of Lobule VII of the cerebellar hemispheres multiplex two different signals when monkeys learn a new arbitrary stimulus–response association. One is the previously described reinforcement learning error signal (Sendhilnathan et al., 2020), and the second, described here, is a signal that describes the state of learning. Neither participates in encoding specific motor kinematics. We suggest that this mixed selectivity facilitates the flexible learning of a new visuomotor associations (Rigotti et al., 2013).

Materials and Methods

We have already described the methods in detail previously (Sendhilnathan et al., 2020). Here we describe it briefly.

Animal subjects

We used 2 male adult rhesus monkeys (*Macaca mulatta*; Monkeys B and S, weighing 10–11 kg each) for the experiments. All experimental protocols were approved by the Animal Care and Use Committees at Columbia University and the New York State Psychiatric Institute and complied with the guidelines established by the Public Health Service Guide for the Care and Use of Laboratory Animals.

Task

We used the National Institutes of Health REX system for behavioral control (Hays et al., 1982). The monkey sat inside a dimly lit recording booth, with its head firmly fixed, in a Crist primate chair 57 mm in front of a back-projection screen on which visual images were projected by a Hitachi CP-X275 LCD projector controlled by a Dell PC running the NIH VEX graphic system.

Two-alternative forced-choice discrimination task. The task began with the monkeys grasping two bars (see Fig. 1A), one with each hand, after which a white $1^\circ \times 1^\circ$ square appeared as a trial cue for 800 ms. Then one of a pair of symbols appeared, briefly for 100 ms in some sessions or until the monkey initiated a hand response in other sessions, at the center of gaze. One symbol signaled the monkey to release the left bar and the other to release the right bar. We rewarded the monkeys with a drop of juice for releasing the hand associated with that symbol.

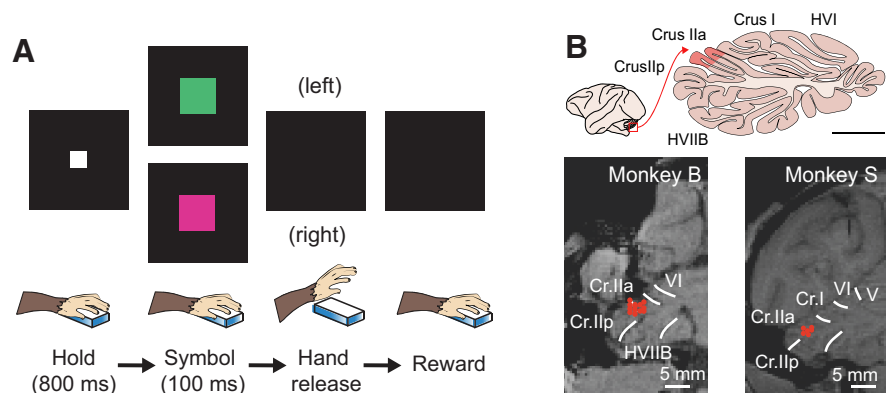


Figure 1. Visuomotor association learning paradigm and cerebellar recording. **A**, Schematic of the two-alternative forced-choice discrimination task. **B**, Top, Schematic of the macaque brain with a sagittal section of the cerebellum showing broad region of the recording location (marked in red). Scale bar, 5 mm. Bottom, The actual recording locations (red markers) superimposed on T1-MRI for both monkeys.

We did not punish the monkeys for errors. The monkeys were trained to only release one hand in response to the presented symbol; if they released both hands, the trial was automatically aborted.

We first trained the monkeys to associate a specific pair of symbols (green square and pink square) with specific hand movements (left- and right-hand release, respectively) for ~4–6 months until their performance was >95% correct; we refer to this as the overtrained condition (OT).

In the visuomotor association learning task, we began each session with the OT symbols; and then after ~30 trials, we switched the symbols to arbitrary fractal symbols which the monkey had never seen before. We refer to this as the N learning condition. The manipulanda the monkeys held on to and released remained the same throughout this task.

In the manipulanda change task, we began every recording session by presenting the monkeys with the same OT symbol pair and bar manipulanda, and after a number of trials, switched the bar manipulanda to dowel manipulanda. The visuomotor association remained the same throughout the task.

Data collection

Single-unit recording. We introduced glass-coated tungsten electrodes with an impedance of 0.8–1.2 MOhms (FHC) into the left mid-lateral cerebellum (near Crus I and II; see Fig. 1B) of monkeys every day that we recorded using a Hitachi microdrive. We passed the raw electrode signal through an FHC Neurocraft head stage, and amplifier, and filtered through a Krohn-Hite filter (bandpass: lowpass 300 Hz to high-pass 10 kHz Butterworth), then through a Micro 1401 system (CED Electronics). We used the NEI REX-VEX system coupled with Spike2 (CED Electronics) for event and neural data acquisition. We verified all recordings offline to ensure that we had isolated P cells and that the spike waveforms had not changed throughout the course of each experiment. We identified cerebellar P cells by the presence of complex spikes online, and offline by the (1) spike waveforms, (2) a pause in simple spike after a complex spike, and (3) the simple spike interspike interval distribution (Dijk et al., 2013).

MRI. After recording the activity of a task-related P cell with an electrode, we removed the microdrive and secured the electrode to the guide tube with a dab of acrylic. We then scanned the monkey in the 3T Siemens MRI scanner using a Kopf MRI compatible stereotaxic apparatus and estimated the lobules from an adjacent slice which did not have the electrode artifact and extrapolated the lobule identity.

Hand tracking. We either painted a spot on the monkeys' right hand with a UV-blacklight reactive paint (Neon Glow Blacklight Body Paint) before every session or tattooed the right hand with a spot of UV Blacklight tattoo ink (Millennium Mom's Nuclear UV Blacklight Tattoo Ink). We used a 5W DC converted UV black light illuminator to shine light on the spot. Then we used a high speed (250 fps) camera (Edmund Optics), mechanically fixed to the primate chair, to capture a video

sequence of the hand movement while the monkeys performed the tasks. We only tracked the monkeys' right-hand movement because the neurons had similar response with either hand movement. We used the track mate feature (Schindelin et al., 2012; Tinevez et al., 2017) and custom-written software in MATLAB to semi-manually track the fluorescent paint spot painted on the monkey's hand.

Data analysis

Criteria for learning. We constructed the learning curve for every session by calculating the percent correct trials in a sliding window of 10 trials shifted by 5 trials. If the monkeys reached >90% correct through the above method and remained >80% for at least the next 20 trials, the associations were considered "learned."

MRD method to detect changes in activity pattern. The change in activity for different P cells occurred at different times in a trial. Given this heterogeneity in its timing, there is no way to predict the time or the magnitude of the activity change, *a priori*, for a randomly selected P cell. Therefore, we wanted to use a method that is blind to the time of change, its distribution, or the intrinsic property of the neuron and more importantly, makes no *a priori* predictions about their properties, in such a way that, given the input (for example, activity in OT and activity in L), it classifies activities as same or different across trial conditions. Since the change in activity could occur at any time of the trial for a given neuron, we used the whole trial activity (from before symbol onset through after reward) for this calculation. However, to avoid the trivial possibility where the two activities are different merely because of changes in reaction time latencies, we limited our analyses window to the following: −400 to 200 ms aligned to symbol onset, and −200 to 600 ms aligned to movement onset for each neuron. The activity aligned to symbol onset and movement onset of the k^{th} trial for each condition (a_k^s and a_k^m , respectively) were concatenated to form single activity vectors for each trial: $a_k = [a_k^s \ a_k^m]^T$. We used these concatenated vectors for all further analyses.

To quantify the change in activity pattern between two conditions (A and B), we computed the mean distance between the two conditions' activity vectors, where each vector was averaged over 10 randomly drawn trials within each condition. Clearly, we first identified the condition with least intracondition variance (suppose Condition A). We then compared the activity within the condition with the least intracondition variance (A) with the activity across both conditions (A and B). To do this, first, we randomly sampled 10 trials each from the last 20 trials in Condition A and the first 20 trials in the Condition B and calculated the root mean squared (rms) distance between the mean of the sampled 10 concatenated vectors (a_k) as follows:

$$rms_{AB} = \sqrt{\sum_{i=1}^n (A_i - B_i)^2}$$

Where $A = \frac{\sum_{i=1}^{10} a_{k_i}}{10}$, $B = \frac{\sum_{i=1}^{10} b_{k_i}}{10}$ and n is the length of A (which is the same as length of B). k_i denotes the index of the i^{th} randomly chosen trial.

We repeated this process 250 times to obtain a distribution of rms distances that compared the extent of change in across-condition activity profile between Conditions A and B. To compare this distribution with the within-condition activity profile, we randomly sampled 10 trials twice without replacement, from Condition A (defined above as the condition with the least intracondition variance) and repeated the same analysis to obtain another distribution of rms distances.

$$rms_{AA} = \sqrt{\sum_{i=1}^n (A_i - \hat{A}_i)^2}$$

Where $A = \frac{\sum_{i=1}^{10} a_{k_i}}{10}$, $\hat{A} = \frac{\sum_{i=1}^{10} \hat{a}_{k_i}}{10}$, and n is the length of A (which is the same as length of $\hat{A}$). k_i denotes the index of the i^{th} randomly chosen trial.

We then calculated the means of each of the two rms distributions, called MRD.

$$MRD_{AB} = \text{mean}(rms_{AB}) \text{ and } MRD_{AA} = \text{mean}(rms_{AA}).$$

Finally, we repeated this process for each P cell and performed a paired t test between the population of MRD_{AB} and MRD_{AA} . We validated our method by running simulated Gaussian signal and Poisson spike trains with different and same variances. We used this method to compare both neural spike density functions and movement trajectories between different conditions.

Analysis of state change epoch. We performed an ANOVA among the simple spike activities in four learning phases: 20 trials in the OT just before the symbol switch, 20 trials just after the symbol switch, 20 trials after 40 trials of symbol switch, and 20 trials after learning. For each learning phase, we considered the activity in two epochs: (1) −1400 to 400 ms from the symbol onset and (2) −400 to 600 ms from the movement onset. We chose this to maximize the sampled trial duration taking the reaction time into account. We chose to extend Epoch 1 until 400 ms after the symbol onset and Epoch 2 from 400 ms before the movement onset to take into account, the long reaction times (800 ms) during early learning trials. After performing an ANOVA in both these epochs, we corrected for multiple comparisons using the Benjamini & Hochberg/Yekutieli false discovery rate control procedure ($Q=0.05$). then, we corrected for multiple comparisons across neurons (p values in the state change epoch) using the Bonferroni correction, and we cross-validated the state change epochs: We sorted the cells on 50% of randomly selected trials and analyzed the data on the held out 50% of the trial, and we confirm that the state change epoch's distribution in time was robust. We defined the start of the state change epoch as the first time point where the corrected p value became significant and remained significant continuously for the next 150 ms. We defined the end of the state change epoch as the first time point when the corrected p value became nonsignificant and stayed nonsignificant continuously for at least the next 250 ms. For the analysis in Figure 7, we included only 75 neurons that exhibited a stable activity throughout the recording. Because this sample was originally chosen to analyze the progression of the reinforcement learning error signal through complete learning sessions (Fig. 5) (Sendhilnathan et al., 2020), all the cells were selective either for success or failure on the prior trial.

Epoch of significant complex spikes activity. We estimated the epochs where the CS had significant activity by performing a t test between the CS activity in 100 ms bins and a baseline activity (−100 to 0 ms aligned to cue onset). Then, we corrected for multiple comparisons using the Benjamini & Hochberg/Yekutieli false discovery rate method.

Statistics

To check whether two independent distributions were significantly different from each other, we first performed a two-sided goodness of fit Lilliefors test, to test for normality, then used an appropriate t test; or else a nonparametric Wilcoxon signed-rank test. All error bars and shading in this study, unless stated otherwise, are mean $\pm$ SEM.

Results

Two monkeys performed a two-alternative forced-choice discrimination task, where the monkeys learned to associate one of two visual symbols with a left-hand movement and the other symbol with a right-hand movement. The monkeys began each trial by placing their hands on each of two manipulanda, after which one of the two symbols appeared on the screen and the monkeys lifted the hand that was arbitrarily associated with that symbol to earn a liquid reward (Fig. 1A; see Materials and Methods).

Visuomotor association learning-related changes in simple spike activity were independent of motor kinematics

We began each experiment with the OT symbols; and then after ~30 trials, we switched the symbols to arbitrary fractal symbols which the monkey had never seen before. We refer to this as the

learning (L) condition. The monkeys performed the OT task with close 100% accuracy. However, once we switched the OT symbols to L symbols, their performance dropped to chance level. The monkeys learned the arbitrarily assigned correct symbol–hand association through trial and error, usually in ~50–70 trials on an average, through an adaptive learning mechanism.

Here, we describe further analysis of the activity of single P cells recorded in Lobule VII near Crus I and II of the cerebellar hemisphere, while monkeys performed the visuomotor association task (Sendhilnathan et al., 2020). During the OT condition, most P cells (106 of 128) significantly increased their firing rate during the bar-release hand movement. Therefore, given the prominent role of the cerebellum in motor control (Manto et al., 2012; Bastein and Leisberger, 2021), if the P cells only encoded the kinematics and the motor error of the hand movements, as long as the hand movement made by the monkeys remained the same, we should expect the neural activity to not change when the monkeys had to learn a new visuomotor association. However, contrary to this, we found that, when the task changed from the OT to the L condition, the activity of 105 of 128 P cells changed dramatically ($p < 0.001$; Kolmogorov–Smirnov test) in several epochs of the trial. Often the cells had significant changes in more than one epoch (Fig. 2). Across the population, the change in activity occurred at different epochs for different neurons (Fig. 2). Furthermore, in some epochs, the activity in the OT condition was higher than that of the learning condition, in other epochs, lower (Fig. 2). Therefore, there was heterogeneity in terms of the number, duration, and the timing of the epochs where these changes were observed and the valence of the change of neural activity in that epoch.

To show that the activity profile between the two conditions was significantly different despite this heterogeneity, in a way that is neither epoch- nor neuron (statistical distribution)-dependent, we computed the rms distance between the spike density functions across the OT and L conditions with repeated random sampling and compared the resulting distribution's mean (mean rms distance, MRD_{across}) with the mean from a distribution of two random samples repeatedly drawn without replacement, from within the OT condition (MRD_{within}), compensating for differences because of reaction times (see Materials and Methods; Fig. 3C, inset). The within-condition MRD values, for the population, were significantly lower than the across-condition MRD values, indicating that the neural activity in the L condition was significantly different from the activity in the OT condition ($p < 0.001$ Wilcoxon signed-rank test; Fig. 3C). To see whether the changes in P-cell activity were accompanied by a change in the kinematics of the monkey's movement, we painted a fluorescent dot on the monkey's hand, and recorded its x - y position with time, using a video camera running at 200 frames/s (see Materials and Methods). Although the neural activity changed dramatically at the symbol switch, the monkeys showed no significant difference in motor kinematics between the OT and L conditions (single session example: Fig. 3D; population: Fig. 3E, $p = 0.38$, paired t test). We did not observe any other significant changes in motor pattern.

Furthermore, the change of neural activity at the symbol switch was unrelated to any change in reaction time at the

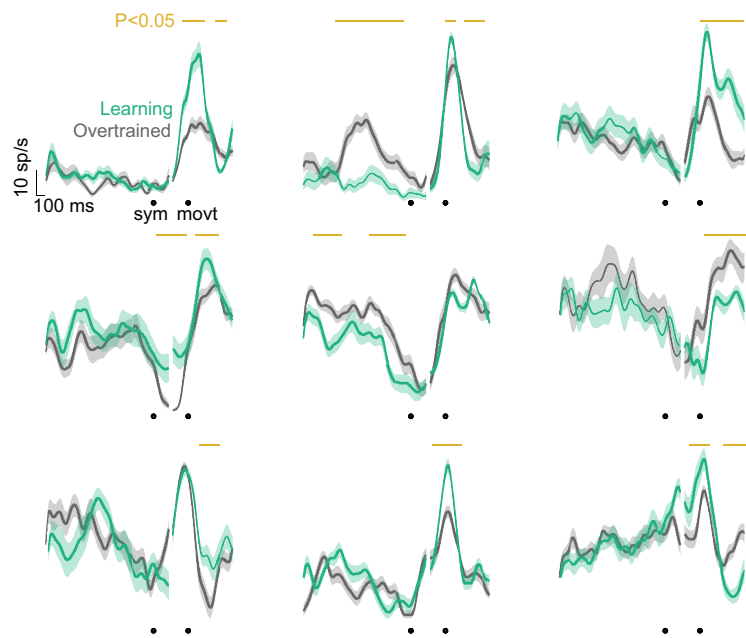


Figure 2. Example single P-cell SS activity in the OT and N learning conditions. Nine example P cells that had differences in neural activities between OT and learning conditions in terms of the timing, duration, number of the epoch of difference, as well as the sign of change in those epochs.

symbol switch: Although in most experiments the reaction time increased at the symbol switch (Fig. 4A), the reaction time did not change at the symbol switch on 24 of 105 (23%) sessions (Fig. 4B), although the monkeys' performance decreased significantly, and the neural activity changed significantly in these sessions (Fig. 4C).

Next, we investigated whether this change in activity profile were merely because of a switch in symbols or whether it was because of the necessity to learn new associations. To do this, we reversed the symbol–hand association once the monkeys had learned the L association. In this symbol reversal paradigm ($N = 25$; Fig. 5A, left), the visual cues that we presented to the monkey remained the same, but the association between the symbols and the hands reversed. That is, the symbol previously cuing a right-hand movement now cued a left-hand movement, and vice versa. Immediately after the reversal, as expected, the performance of the monkeys dropped below chance level and the monkeys took a longer time to learn the association (Fig. 5A, right). Again, in this condition, the activity profile of the P cells changed significantly ($p < 0.001$, t test; Fig. 5B), implying that the change in neuronal activity was dependent on the monkey's having to learn a new association, but was not related to the changes in the symbols per se.

Although the visuomotor reinforcement learning task was not associated with any change in movement kinematics, we then investigated whether forcing a change in kinematics by changing the manipulanda the monkey used to report its decision would be accompanied by a change in the activity of the P cells, although the symbols and their movement association did not change. This experiment began with the monkeys performing the usual OT task, starting the task by placing their hands on two bar manipulanda, and reporting their decision by lifting one of their hands. Then, we changed the motor aspects of the task, having the monkeys grab dowel manipulanda (Fig. 6A) and report their decision by releasing these dowels, which involved a much different movement. The kinematics differed significantly between the two versions of the task (Fig. 6B),

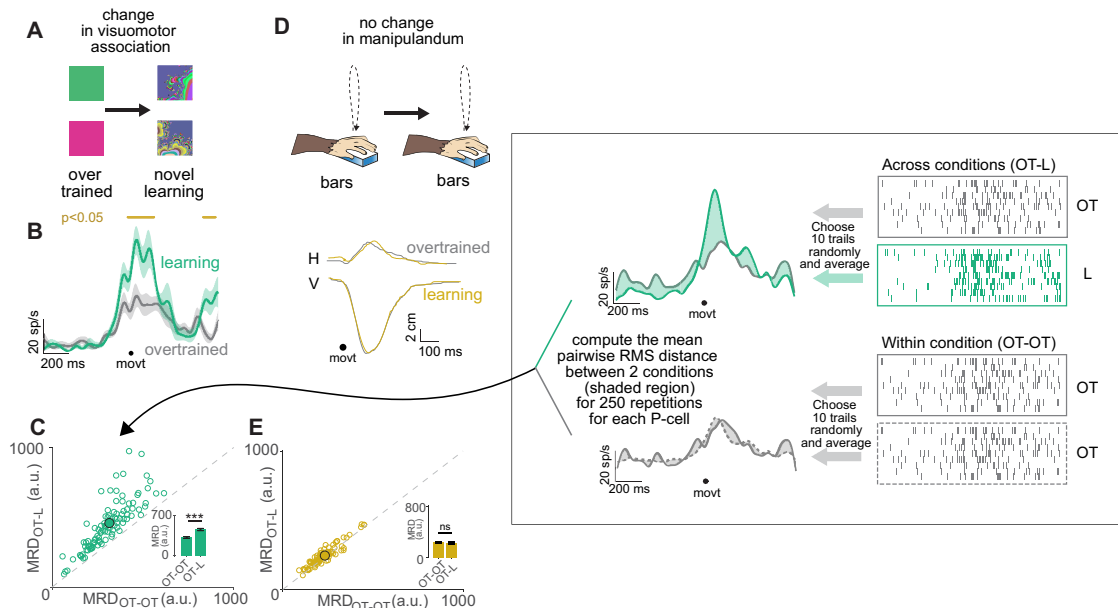


Figure 3. P-cell activity distinguished OT and N symbols despite the absence of changes in movement kinematics. **A**, Task in which there was a change in visuomotor association from OT to N association. **B**, Spike density plot of a representative P cell aligned to bar-release hand movement onset (movt) that responded to the symbol switch by firing differently between the OT (gray) and N (green) conditions. Gold line on top indicates the epochs where this difference was significant ($p < 0.05$, paired t test). **C**, Scatter plot of MRD for within OT condition (OT-OT) versus MRD for across OT and L learning (L) condition (OT-L) (see Materials and Methods). Open circle represents a P cell ($N = 105$). Filled circle represents the mean value. Inset, Mean MRD for OT-N and OT-OT conditions. $***p < 0.001$ (Wilcoxon signed-rank test). Inset, Top, Schematic of the MRD method. First, we randomly sampled 10 trials each from the last 20 trials in the OT condition (gray) and the first 20 trials in the L condition (green) and calculated the rms distance between the mean activities. We repeated this process 250 times to obtain a distribution of rms distances that compared the extent of change in across-condition activity profile in the L condition from the activity profile in the OT condition. Bottom, To compare this distribution with a control null distribution, we randomly sampled 10 trials twice without replacement from the OT condition and repeated the same analysis to obtain another distribution of rms distances to obtain an estimate of variability of within-condition. A test of statistical significance between the mean of these two distributions would provide an estimate of the change between conditions as shown in Figure 3C. **D**, Diagram bar-release hand movement trajectories (top) and the actual movement trajectories decomposed into horizontal (H) and vertical (V) components traces (bottom) for OT (gray) and N (yellow) conditions. **E**, Same as in **C**, but for bar-release hand movement trajectories. Not significant, $p = 0.3822$ (paired t test).

although symbols were the same and had the same visuomotor association (Fig. 6C). The P-cell response did not change when the kinematics changed (example single cell: Fig. 6D; population: $p = 0.23$; paired t test; $N = 31$; Fig. 6E).

The P cells changed their activity state as the monkeys learned the task

Next, we studied how the transient change in neural activity profile from OT activity to the activity at the beginning of learning (Figs. 2 and 3) changed through learning. The P cells changed their activities with learning in three ways: First, 63% of the P cells that showed an increase in firing rate from the OT to beginning of learning continued to increase their firing rate through learning, showing a “positive state change” (for single-neuron example, see Fig. 7A; see Fig. 7D for population). The remaining 37% of the neurons eventually returned to the activity, comparable to the OT condition (for single-neuron example, see Fig. 7B, top; see Fig. 7D for population), showing “no permanent state change.” Of the P cells that showed a reduction in firing rate from the OT to beginning of learning, 70% of neurons continued to decrease their firing rate through learning, showing a “negative state change” (for single-neuron example, see Fig. 7C; see Fig. 7D for population) while the remaining 30% of the neurons eventually returned to an activity profile comparable to the OT condition (for single-neuron example, see Fig. 7B, bottom; see Fig. 7D for population) again showing “no permanent state change.” The neurons did not change the time of their peak activity relative to the movement. However, the epoch of the state change occurred at different times in the trial for different neurons (Fig. 7A–C). The epoch of the state change during the trial

was not necessarily the time of peak activity during the trial (Fig. 7B, bottom, C). Across the population, all three types of neurons were approximately equally distributed.

Complex spikes did not instruct changes in simple spike neural activity during learning

The conventional model of the cerebellum (Marr, 1969; Albus, 1971; Ito, 1984) posits that the complex spikes, driven by the climbing fiber input from the inferior olive, provide an error signal that affects the sensitivity of the Purkinje cell simple spikes to the input of the mossy fibers as transmitted by the granule cells. We asked whether the complex spike activity (Fig. 8A,B) induced these learning-related changes in simple spike neural activity. We found no relationship between the time of significant modulation of complex spike activity (see Materials and Methods) and the time of the epoch of state change for a given P cell (Fig. 8C, D). Furthermore, the simple spike activity was similar in trials whose previous trials did or did not have complex spikes 175–50 ms before the state change epoch. (Fig. 8E). This suggests that there is no obvious relationship between simple spike and complex spike activity during reward-based learning (Larry et al., 2019; Sendhilnathan et al., 2021). and that the complex spike responses did not affect the simple spike activity or the behavior through an error-based learning mechanism.

Discussion

We recorded the simple spike and complex spike activities of P cells in the Crus I and Crus II of the cerebellar hemispheres while monkeys performed a familiar visuomotor association and while

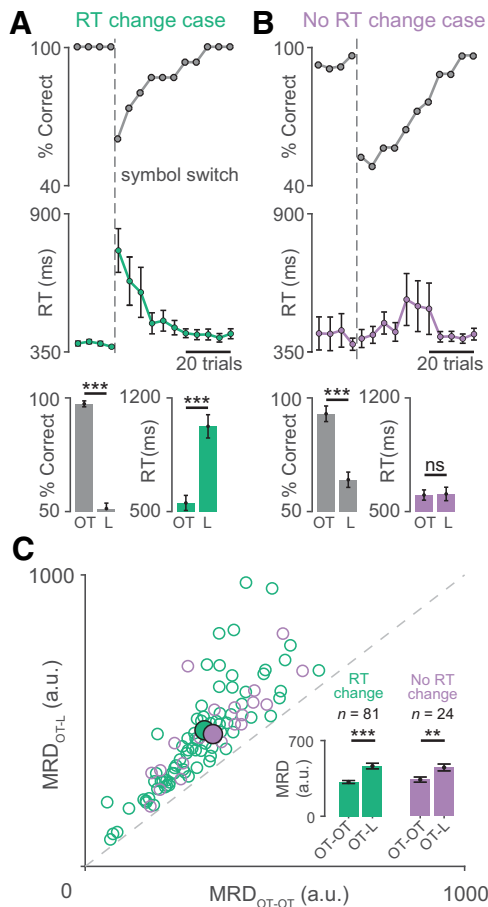


Figure 4. Neural activity changes were independent of changes in reaction time. **A**, Top, Percent of correct trials plotted as a function of trial number relative to the switch to N visuomotor association. Middle, Reaction times for the same trials. Error bars indicate SEM. Bottom, Mean percent correct (left) and reaction time (right) in the OT and L learning (L) conditions for all the sessions with changes in reaction time. $***p < 0.001$ (Wilcoxon rank sum test). **B**, Percent correct (top), reaction time (middle), and session averages (bottom) for sessions in which the manual reaction time did not change (not significant, $p = 0.8151$, Wilcoxon rank sum test) after the switch to N visuomotor association but the performance did ($***p < 0.001$, Wilcoxon signed-rank test). Same convention as in **A**. **C**, Same data as in Figure 3C, but separated into sessions with RT change condition ($N = 81$; green) ($***p < 0.001$, Wilcoxon signed-rank test) and no RT change condition ($N = 24$; violet; $**p < 0.001$, paired t test). MRD values indicating that the change in neural activity in the L condition was significantly different from the activity in the OT condition ($p < 0.001$ Wilcoxon signed-rank test; Fig. 3C) regardless of whether this was accompanied by a change in RT.

they learned a new visuomotor association. We noted several key differences between the neural signatures in these two conditions. First, different neurons showed activity changes in different epochs of the trial at the symbol switch. This change in neural activity was neither because of changes in reaction time or movement kinematics, nor was it because of changes in visual context since the reversal learning followed a similar trend as well. We found that only in the cases that required reinforcement learning of new visuomotor associations did the P cells change their neural activity to reflect that learning (Table 1).

During learning, induced by the symbol switch, the population of P cells modulated their response in broadly three different ways. Compared with their neural activity in the familiar task, some neurons kept increasing (positive state) or decreasing (negative state) their firing rate as the animal learned the task; hence, their activity after learning the new task was significantly different from the activity in the familiar task. A few other neurons,

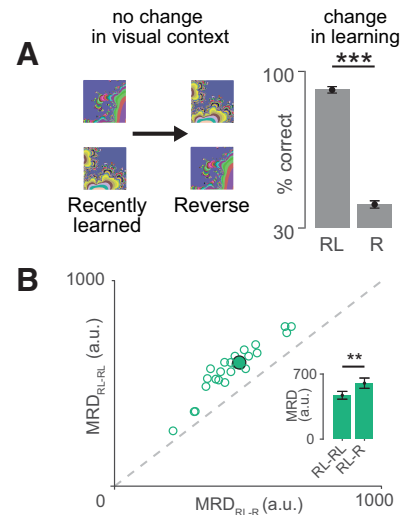


Figure 5. P cells changed their activity profile in the symbol reversal task. **A**, Left, Changing the visuomotor association of previously learned symbols from recently well-learned (RL) to reversal learning (R) condition but with new learning. Right, Mean behavioral performance in the recently learned (RL) and the reversal (R) conditions. $***p < 0.001$ (Wilcoxon signed-rank test). **B**, Scatter plot of MRD for within recently learned condition (RL-RL) versus MRD for across recently learned and reversal condition (RL-R). Same convention as Figure 3C. $N = 25$ P cells. $**p < 0.01$ (paired t test).

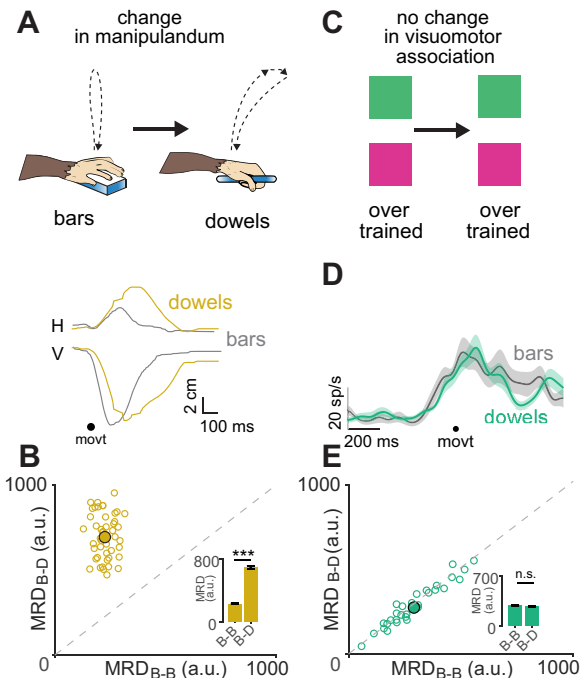


Figure 6. P cells did not respond to a change in motor kinematics in the absence of a symbol change. **A**, Top, Diagram represents different hand movement trajectories with the change in manipulanda from bars (B) to dowels (D). Bottom, Actual movement trajectories decomposed into horizontal (H) and vertical (V) component traces for bars (gray) and dowels (yellow) conditions. **B**, Same as in Figure 3E, but for hand movement trajectories in bars-dowels condition. $***p < 0.001$ (paired t test). **C**, Task in which the visuomotor association did not change. **D**, Same representative neuron from Figure 3B when the movement changed but the association did not. **E**, Same as in Figure 3C, but for neural activity in bars-dowels condition ($N = 31$). Not significant, $p = 0.23$ (paired t test).

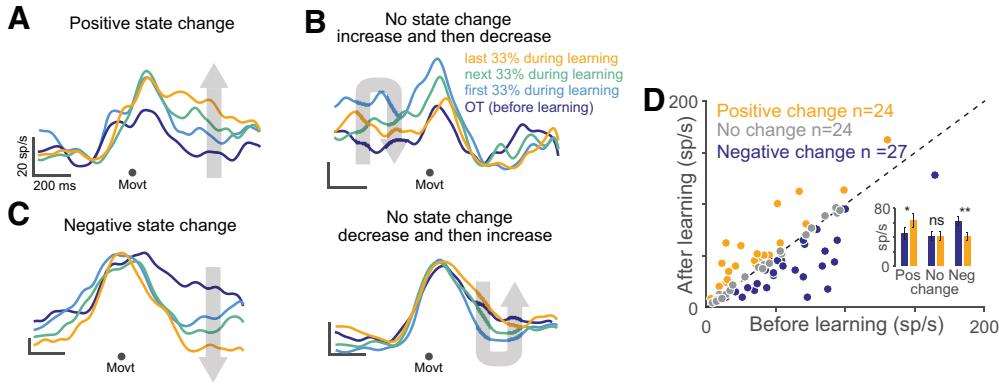


Figure 7. Learning-dependent state changes in neural activity. **A**, Representative neuron whose activity increases at the symbol switch and also increases through learning, showing a positive state change. Activity was synchronized on the movement (dot). **B**, Top, Representative neuron whose activity increases at the symbol switch but decreases through learning, returning to initial state, thus showing no state change. Bottom, Representative neuron whose activity decreases at the symbol switch but increases through learning, returning to initial state, thus showing no state change. **C**, Representative neuron whose activity decreases at the symbol switch and also decreases through learning, showing a negative state change. **D**, Activity for trials before learning and after learning for all cells. Each dot is a cell. Positive state change P cells ($N = 24$) lay above the diagonal. Negative state change P cells ($N = 27$) lay below the diagonal. No state change P cells ($N = 24$) lay on the diagonal. Inset, The mean firing rate for all three classes of neurons before and after learning. $*p < 0.05$, $**p < 0.01$, t test.

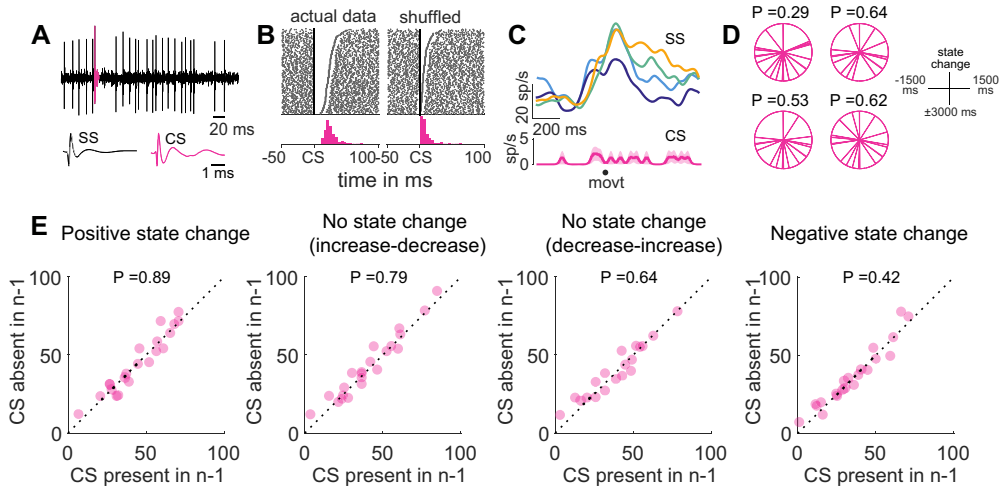


Figure 8. Complex spikes did not instruct simple spikes during reinforcement learning. **A**, Top, Representative neural recording with simple spikes (gray) and complex spike (pink). Bottom: typical waveforms of a simple spike (SS) and a complex spike (CS). **B**, Time of simple spike occurrence relative to complex spike occurrence. Top: Each dot is the time of a simple spike. The raster lines are synchronized on the time of a complex spike (CS, time = 0) and ordered from top to bottom by the interval between the complex spike and simple spike. Bottom: Complex spike triggered simple spikes for the actual data (i.e., $\Pr(SS = t | CS = 0)$ where t is the time of occurrence of an SS; left) and shuffled (i.e., $\Pr(SS = t | CS = x)$ where t and x are the times of occurrences of an SS and a CS, respectively; right). **C**, Top, simple spike (SS) response from the top cell in Figure 7A with complex spike (CS) response (bottom). Bottom: Distribution histograms of the interval between the complex spike and the next simple spike. Left, actual data. Right, shuffled data. **D**, Polar plot (of the entire trial period) of time of significant complex spike modulation during learning relative to the time of state change for each cell. The time of the state change is at the top of each time circle. The intersection of each line with the time circle represents the interval between time of significant modulation of complex spikes (in ms) relative to the time of state change (see insert). Top left: positive state change (Figure 7A). Top right: negative state change (Figure 7C). Bottom right: no state change with increase and then decrease (Figure 7B top). Bottom left: no state change with decrease then increase (Figure 7B bottom). P values were from a circular Rayleigh z test. **E**, Simple spike activity with a complex spike present in the previous trial (abscissa) versus simple spike activity on trials with a complex spike absent in the previous trial (ordinate) for different types of neurons. P values were from a paired t test.

Table 1. Summary of results

Experiment	Change in sensory information	Change in movement	Change in reaction time	Change in reinforcement learning	Change in P-cell activity
Symbol switch	✓		✓	✓	✓
Symbol reversal			✓	✓	✓
Manipulanda switch		✓			

however, showed an initial change in activity after the symbol switch, but their activity after learning returned to a level similar to the activity in the familiar task. Thus, the population of P cells describes the state of learning without specifying the details of the monkey's behavior (Fig. 7).

It could be possible that the change in neural activity immediately after the symbol switch could be a result of attention or arousal. Spatial attention is the selection of a spatial location for further processing, often accompanied by a saccade to that location. The monkeys usually begin to look at

the fixation point (Sendhilnathan et al., 2020) which precedes the appearance of the symbol, yet most cells have their greatest activity around the movement even when the change evoked by the symbol switch does not occur around the movement. Furthermore, spatial attention is accompanied by a decrease in reaction time (Posner, 1980), yet in most of our experiments, the reaction time was greatest at the symbol switch and decreased with learning, and on other experiments the reaction time did not change, and the P cells changed their neural activity regardless of this change in reaction time at the symbol switch (Fig. 4). Feature attention is the enhancement of response to a selected feature regardless of the locus a spatial attention. However, because the symbol only remains on the screen for 100 ms, feature location is irrelevant. Arousal is generally associated with an increase in neural activity (Zhang et al., 2014). We often see the same neurons that show both increased and decreased activity because of the symbol switch (Fig. 2). Furthermore, some neurons only showed either an increase or a decrease in activity (Fig. 2).

We previously described a reinforcement error signal (Sendhilnathan et al., 2020) carried by the same P cells analyzed in this study, that is greatest at the symbol switch and approached zero as the monkeys learned the task, disappeared in the OT task. Although the magnitude of the reinforcement error signal decreased with learning for all the P cells, here we show that only some neurons returned their activity (~33%; Fig. 7B) to their initial state, while the rest had significantly different activities between the end of learning and the OT conditions (Fig. 7A,C). Therefore, for these neurons, although they no longer report an error signal after learning, their neural activity gets stabilized at a different state because of learning-related changes. This was similar to previous observations in motor learning paradigms in other studies (Medina and Lisberger, 2008).

A hallmark of the previously reported reward-based error signal was that it was unlikely to be driven by complex spike activity (Sendhilnathan et al., 2020, 2021). Here, too, the concurrent complex spike activity was unlikely to have caused these changes in simple spike activity (Fig. 8) or could have been the reason why different P cells had different learning-related activities. This is because we did not find any correlation between the time of simple spike activity and the time of complex spike activity. These results are in line with the previous studies which showed a learning dependency on the activity of simple spike, independent of the complex spike in reward-based learning (Larry et al., 2019) and motor learning (Ke et al., 2009; Streng et al., 2018; Avila et al., 2021).

Since most of the neurons increased the activity before or during the movement, one can argue that these neurons simply encoded the motor command for hand movement, and that the modulation we found depended on the change of the movement kinematic from a familiar to a N task. However, we can rule out this hypothesis since we did not find any change in the kinematics of the hand movement made by the monkey after the symbol switch (Fig. 3). Furthermore, when we induced a change in hand movement through a manipulandum switch, despite changes in hand movement kinematics, the P-cell neural activity remained unchanged (Fig. 6). This suggests that the P cells did not encode the exact kinematics of the hand movement but, rather, goal of the action.

The area of the cerebellum that we have studied lies within a closed loop network involving the PFC and the basal ganglia (Bostan et al., 2013). Neurons in these areas are

active during visuomotor reinforcement learning, with activity related to reward and movement, although the details of this activity differ from the activity we demonstrated here. For example, in the PFC (Asaad et al., 1998), neurons specify the movement that the monkey will make, and learning is manifest by a lengthening of the interval between the neuronal selection of the movement and the actual movement. Conversely, here we show that, although the activity of neurons changes during learning, the neurons do not specify the actual movement, nor does the time of their peak activity change relative to the movement. Some prefrontal neurons distinguish between prior reward and prior failure (Histed et al., 2009), but this reward-related activity does not seem to change during learning. Both areas show signals with mixed selectivity: the cortex multiplexing selectivity for the symbol and the movement (Rigotti et al., 2013), the cerebellum multiplexing signals for learning error and state. Nonetheless, reinforcement learning is impaired by inactivation of PFC (Murray et al., 2000) as well as the region of the cerebellum whose activity is described in this report (Sendhilnathan and Goldberg, 2020). Although neuronal activity with only one form of selectivity is simple to understand, mixed selectivity intuitively provides challenge. However, Rigotti et al. (2013) have shown that the population activity of mixed selectivity neurons can be decoded to provide independent measures of each of the multiplexed signals. This suggests that the mixed selectivity we have demonstrated here may provide a more efficient way for the brain to code complex and flexible learning behaviors.

References

- Albus JS (1971) A theory of cerebellar function. *Math Biosci* 10:25–61.
- Asaad WF, Rainer G, Miller EK (1998) Neural activity in the primate prefrontal cortex during associative learning. *Neuron* 21:1399–1407.
- Avila E, Flierman NA, Holland PJ, Roelfsema PR, Frens MA, Badura A, De Zeeuw CI (2021) Purkinje cell activity during suppression of voluntary eye movements in rhesus macaques. *bioRxiv*. Doi:10.1101/2021.03.26.437236.
- Bastian A, Lisberger SG (2021) Cerebellum. In: *Principles of neural science*, sixth edition (Kandel EK, Koester JD, Mack SH, Siegelbaum SA), pp 908–931. New York: McGraw-Hill.
- Bostan AC, Dum RP, Strick PL (2013) Cerebellar networks with the cerebral cortex and basal ganglia. *Trends Cogn Sci* 17:241–254.
- Caligiore D, Pezzulo G, Baldassarre G, Bostan AC, Strick PL, Doya K, Helmich RC, Dirks M, Houk J, Jörntell H, Lago-Rodriguez A, Galea JM, Miall RC, Popa T, Kishore A, Verschure PF, Zucca R, Herreros I (2017) Consensus paper. Towards a systems-level view of cerebellar function: the interplay between cerebellum, basal ganglia, and cortex. *Cerebellum* 16:203–229.
- De Zeeuw CI, Ten Brinke MM (2015) Motor learning and the cerebellum. *Cold Spring Harb Perspect Biol* 7:a021683.
- Dijk G, Hulle MM, Heiney SA, Blazquez PM, Meng H, Angelaki DE, Arenz A, Margrie TW, Mostofi A, Edgley S (2013) Probabilistic identification of cerebellar cortical neurones across species. *PLoS One* 8:e57669.
- Giovannucci A, Badura A, Deverett B, Najafi F, Pereira TD, Gao Z, Ozden I, Kloth AD, Pnevmatikakis E, Paninski L, De Zeeuw CI, Medina JF, Wang SS (2017) Cerebellar granule cells acquire a widespread predictive feedback signal during motor learning. *Nat Neurosci* 20:727–734.
- Hays AV, Richmond BJ, Optican LM (1982) A UNIX-based multiple process system for real-time data acquisition and control. *WESCON Conf Proc* 2:1–10.
- Heffley W, Hull C (2019) Classical conditioning drives learned reward prediction signals in climbing fibers across the lateral cerebellum. *Elife* 8:e46764.
- 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.

- Histed MH, Pasupathy A, Miller EK (2009) Learning substrates in the primate prefrontal cortex and striatum: sustained activity related to successful actions. *Neuron* 63:244–253.
- Ito M (1984) *The cerebellum and neural control*. New York: Raven.
- Ke MC, Guo CC, Raymond JL (2009) Elimination of climbing fiber instructive signals during motor learning. *Nat Neurosci* 12:1171–1179.
- Kostadinov D, Beau M, Pozo M, Häusser M (2019) Predictive and reactive reward signals conveyed by climbing fiber inputs to cerebellar Purkinje cells. *Nat Neurosci* 22:950–962.
- Koziol LF, Budding D, Andreasen N, D'Arrigo S, Bulgheroni S, Imamizu H, Ito M, Manto M, Marvel C, Parker K, Pezzulo G, Ramnani N, Riva D, Schmähmann J, Vandervort L, Yamazaki T (2014) Consensus paper. The cerebellum's role in movement and cognition. *Cerebellum* 13:151–177.
- Larry N, Yarkoni M, Lixenberg A, Joshua M (2019) Cerebellar climbing fibers encode expected reward size. *Elife* 8:e46870.
- Manto M, Bower JM, Conforto AB, Delgado-García JM, da Guarda SN, Gerwig M, Habas C, Hagura N, Ivry RB, Mariën P, Molinari M, Naito E, Nowak DA, Oulad Ben Taib N, Pelisson D, Tesche CD, Tilikete C, Timmann D (2012) Consensus paper. Roles of the cerebellum in motor control: the diversity of ideas on cerebellar involvement in movement. *Cerebellum* 11:457–487.
- Marr D (1969) A theory of cerebellar cortex. *J Physiol* 202:437–470.
- Medina JF, Lisberger SG (2008) Links from complex spikes to local plasticity and motor learning in the cerebellum of awake-behaving monkeys. *Nat Neurosci* 11:1185–1192.
- Murray EA, Bussey TJ, Wise SP (2000) Role of prefrontal cortex in a network for arbitrary visuomotor mapping. *Exp Brain Res* 133:114–129.
- Posner MI (1980) Orienting of attention. *Q J Exp Psychol* 32:3–25.
- Raymond JL, Medina JF (2018) Computational principles of supervised learning in the cerebellum. *Annu Rev Neurosci* 41:233–253.
- Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, Fusi S (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497:585–590.
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9:676–682.
- Sendhilnathan N, Goldberg ME (2020) The mid-lateral cerebellum is necessary for reinforcement learning. *bioRxiv*. Doi:10.1101/2020.03.20.000190.
- Sendhilnathan N, Ipata AE, Goldberg ME (2020) Neural correlates of reinforcement learning in midlateral cerebellum. *Neuron* 106:188–198.e5.
- Sendhilnathan N, Ipata A, Goldberg ME (2021) Mid-lateral cerebellar complex spikes encode multiple independent reward-related signals during reinforcement learning. *Nat Commun* 12:1–10.
- Streng ML, Popa LS, Ebner TJ (2018) Complex spike wars: a new hope. *Cerebellum* 17:735–746.
- Tedesco AM, Chiricozzi FR, Clausi S, Lupo M, Molinari M, Leggio MG (2011) The cerebellar cognitive profile. *Brain* 134:3672–3686.
- Tinevez JY, Perry N, Schindelin J, Hoopes GM, Reynolds GD, Laplantine E, Bednarek SY, Shorte SL, Eliceiri KW (2017) TrackMate: an open and extensible platform for single-particle tracking. *Methods* 115:80–90.
- Wagner MJ, Kim T, Savall J, Schnitzer MJ, Luo L (2017) Cerebellar granule cells encode the expectation of reward. *Nature* 544:96–100.
- Zhang M, Wang X, Goldberg ME (2014) A spatially nonselective baseline signal in parietal cortex reflects the probability of a monkey's success on the current trial. *Proc Natl Acad Sci USA* 111:8967–8972.