The basal ganglia achieve a higher signal-to-noise ratio than the cerebellum by sharpening reward and movement signals

Noga Larry[1*], Gil Zur[1*] and Mati Joshua[1]

1. Edmond and Lily Safra Center for Brain Sciences, the Hebrew University, Jerusalem, Israel
   * These authors contributed equally.

The basal ganglia and the cerebellum are major subcortical structures in the motor system. The basal ganglia have been cast as the reward center of the motor system, whereas the cerebellum has been hypothesized to be involved in adjusting sensorimotor parameters. Recent findings of reward signals in the cerebellum have challenged this dichotomous view. To directly compare the basal ganglia and the cerebellum we recorded from oculomotor regions in both structures from the same monkeys. We partitioned the trial-by-trial variability of the neurons into reward and eye-movement signals to compare the coding within and across structures. Reward expectation and movement signals were the most pronounced in the output structure of the basal ganglia, intermediate in the cerebellum, and the smallest in the input structure of the basal ganglia. These results indicate different information convergence properties within these structures and demonstrate differences in internal computations.

Introduction

To produce behavior, the brain transforms sensory information along with predictions about the consequences of actions into movements. This process is dispersed across multiple brain structures. Thus, to understand how the brain operates, the contribution of each area to behavior production needs to be disentangled. One possible route toward this goal is to constrain the computations of each area by identifying area-specific signals. However, in a complex interconnected system such as the brain, specificity is rare, and signals are often found ubiquitously[1–3]. A system approach that compares signals across populations is thus better positioned to yield insights into the transformations of signals as they propagate across neural populations.

The basal ganglia and cerebellum are major subcortical structures that control movement. The basal ganglia are considered to be involved in choices between possible actions and in performing reward-based learning[4,5]. By contrast, the cerebellum has been hypothesized to be involved in adjusting movement and performing error-based motor learning[4,6,7]. However, recent findings of reward signals in the cerebellum have challenged this dichotomous separation of roles[1,8–12], suggesting that the fundamental difference between these structures may not be so straightforward[13]. Although reward signals in the basal ganglia and cerebellum are similar, they currently cannot be directly compared, since experiments have been conducted on different behaviors, tasks, and species. To overcome these shortcomings, we directly compared signals in the basal ganglia and the cerebellum by recording from oculomotor areas in both structures in the same tasks and monkeys. The monkeys performed eye-movement tasks in which the reward probability was manipulated.

Along the motor system, the neural activity becomes directly related to movement production, as sensory and contextual signals are filtered out[14]. Thus, we would expect reward signals to be present upstream and be transformed into signals that are directly related to movement production closer to the motor neurons. A hierarchical organization of a transformation from reward to movement predicts that reward signals will decrease close to the motor output, whereas movement signals will increase. The cerebellum is anatomically closer than the basal ganglia to the motor neurons that control eye movements[15]. Functionally, electrical stimulations of the oculomotor cerebellum produce eye movements[16,17], whereas stimulation of the oculomotor basal ganglia produces eye movement in a context-dependent manner[18], consistent with the less direct route to the motor neurons. Hence a hierarchical organization model predicts that the reward signals should be more pronounced in the basal ganglia than in the cerebellum and that movement signals should be more pronounced in the cerebellum than in the basal ganglia. Similarly, within structures, we would expect a similar trend in which reward signals are replaced by motor signals as information moves downstream.

However, in sharp contrast with the hierarchical organization assumption, we found that both reward expectation and movement signals were most pronounced in the output structure of the basal ganglia. To study how eye movements and reward signals build along each structure we compared the input or intermediate states of processing with the output neurons. In the basal ganglia, we found a strong sharpening of signals from the input to output whereas in the cerebellum we found that output neurons and local inputs to output neurons had similar signals. Thus, these results highlight a functional difference between transformations in the basal ganglia and cerebellum. In the basal ganglia, relevant information is sharpened, akin to filtering out task-irrelevant noise. In the cerebellum, the lack of

sharpening suggests a convergence of opponent signals on output neurons, which might be beneficial for learning.

Results

Eye movement tasks that manipulate reward probability and movement direction

To compare the coding of eye movements and reward directly in the cerebellum and the basal ganglia, we recorded from both areas in the same monkeys. The monkeys performed smooth pursuit and saccade tasks in which we manipulated the probability of receiving a reward and the direction of eye movements. The trials began with the appearance of a white spot in the center of the screen (Fig. 1A). The monkeys were required to fixate on the spot which then changed into one of two colors, indicating the probability of the reward that would be delivered upon successful completion of the trial. One color corresponded to a probability of 0.75 (blue in Fig. 1A) and the other to a probability of 0.25 (red). In the pursuit task, after a variable delay, the colored target began to move 20°/s in one of eight directions[19] and the monkey was required to smoothly track the target (Fig.1B top, Extended Data Fig. 1A). On the saccade task, the target disappeared and reappeared at one of eight 10° eccentric locations. Shortly (219±44 and 194±37 ms for Monkeys A and G) after the target jumped, the monkey was required to saccade towards the target (Fig. 1B bottom, Extended Data Fig. 1B). At the end of a successful trial, the monkey received or did not receive a reward, at the probability specified by the color of the target.

Before recording the neural activity, we studied the relationship between reward probability and eye movements (see Methods). The 0.25 and 0.75 probabilities were chosen since they generated substantial reward expectation differences that were clearly noticed by the monkeys (Fig. 1C) while keeping the behavioral effects very small (Extended data Fig. 1, Methods section), thereby disambiguating motor and reward signals.

We segmented both tasks into three separate epochs. The first epoch was the cue where information about the probability of future reward was first made available. The second was the motion epoch, in which the target moved, and pursuit or saccade movements occurred. The third epoch was the outcome where the reward was either delivered or omitted.

Directly comparing neurons in the basal ganglia and cerebellum

To study how reward and movements are represented within and between structures, we recorded from areas in the basal ganglia and cerebellum that are involved in the generation of pursuit and saccades[20–25]. In the basal ganglia we recorded from the input structure, the body of the caudate, and the output structure, the substantia nigra pars reticulata (SNpr). In the cerebellum, we recorded from the oculomotor vermis. We identified Purkinje cells, the sole output of the cerebellar cortex, based on the presence of complex spikes followed by a pause in the activity of simple spikes[26,27]. Other neurons were classified as putative local neurons that do not project outside the cerebellar cortex. The cue and outcome epochs of the pursuit and saccade tasks were identical, and the results were similar on both tasks. Thus, when analyzing these epochs we pooled neurons across tasks.

In each population, we found neurons that responded to the presentation of the reward probability cue (Fig. 1D-G). To compare the responses to the cue between the populations, we calculated the reward probability effect size. Specifically, we used the partial ω^2 (ω_p^2) to measure effect size (see Methods). ω_p^2 is an unbiased estimator of the variance explained by an experimental variable, normalized by the same variance plus a noise term[28] (Fig. 1H and I). The noise term is an estimator of the variance which is unaccounted for by any variable in the experiment; i.e., the trial-by-trial variability within task conditions. ω_p^2 values close to 1 indicate large differences between the levels of a variable relative to the noise, and values close to 0 indicate smaller or non-existent differences. For example, the example SNpr neuron (Fig. 1F) had consistent responses within conditions, either pausing or increasing its rate in the low and high probability conditions. In contrast, the caudate, cerebellar Purkinje cell, and the local cerebellar neuron (Fig. 1D, E, and G) showed less consistent responses. These examples demonstrate that the effect size reflects the signal-to-noise ratio in the activity.

This approach has several advantages. First, ω_p^2 is agnostic to the direction of response. Some neurons in our sample had higher rates when reward probability was high (for example, Fig. 1E and F) while others had higher rates when reward probability was low (Fig. 1D and G). All the populations had neurons with higher rates for either the high or the low probability conditions[29,30]. Second, the variance that can be explained by additional variables is removed from the noise term. This is critical for drawing comparisons between populations that might multiplex the coding of experimental variables to different degrees. Finally, since ω_p^2 is an unbiased estimator, when a variable does not affect the firing rate, ω_p^2 will be 0 on average. This contrasts with other effect size measures in which the size of the bias depends on the number of conditions (Fig. 1I).

Fig. 1J shows the reward probability effect size in 50 ms time bins for each of the example neurons. All neurons showed an increase in the effect size following the appearance of the cue when reward probability information first became available. However, this increase was substantially larger in the SNpr neuron. This increase corresponds to the time within the trial in which the difference between the probability condition in the SNpr neuron is large and the inter-condition variability is small. The consistent but different responses to the reward probability condition result in a high effect size. Intuitively, a large effect size will be correlated with improved decoding of the variable (Extended Data Fig. 2A-C).

We also calculated the effect size over the entire trial (see Methods). This allowed us to obtain a single assessment of the reward probability coding through the entire epoch. These whole epoch effect sizes are specified in the caption to Fig. 1D-G. We verified that the effect sizes were robust to different bin sizes and that they were not determined by the firing rates of neurons (Extended Data Fig. 2D-I and Extended Data Fig. 3, see Methods).

<u>The coding of reward probability is more pronounced in the SNpr than in the cerebellum and the caudate</u>

To estimate the coding of reward probability in the populations around the time of cue presentation, we calculated the reward probability effect size of individual neurons and the average across the population (Fig. 2A). We pooled data across both tasks since the

structure of the task in this epoch was identical and the results were the same for the two tasks. As expected from an unbiased estimator before reward probability information was made available, the reward effect size of all populations was close to zero (Fig. 2A, before time 0). After cue presentation, we found a transient increase in the reward probability effect size in all populations. This effect was the most pronounced in the SNpr (Fig. 2C, Permutation Welch's ANOVA test: $p<0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}purkinle}$, $p_{SNpr\text{-}local\ cerebellar}<0.001$). It is unlikely that this result was due to the coding of conditioned licking in the cue epoch because these were similar across the different reward probability conditions (Extended Data Fig. 4).

The increase in reward probability effect size was transient and decreased after the initial increase, although it remained slightly larger than zero in all four populations. When the targets moved, the reward probability effect size remained small in all populations (Fig. 2B and D). SNpr neurons showed a relatively late increase in the coding of reward probability which was not observed in any of the other populations (Fig. 2B and D).

Our analysis thus confirms recent findings of reward-related activity in the cerebellum[1,9–11] and many studies showing reward-related activity in the basal ganglia[30–32]. Further, the SNpr responses were substantially stronger than other populations. The results indicate a strong sharpening of reward signals within the basal ganglia, from the caudate to the SNpr. However, the finding that the caudate exhibited a smaller average effect size than both cerebellar populations does not support a simple hierarchy in the representation of reward probability across structures.

<u>Coding of eye movement direction is more pronounced in the SNpr than in the cerebellum and the caudate</u>

To study how eye movement signals are organized across the subcortical networks we next compared the coding of eye movement directions between the populations in the pursuit and saccade tasks. Some neurons in our sample had complex dynamics in time (for example, Extended Data Fig. 5C). The effect size analysis allowed us to compare the coding of directions between tasks and populations without averaging in time to calculate the tuning curve[30] or having to define the preferred direction of neurons[33]. We therefore calculated the direction effect size of each population and task. All the populations showed an increase in direction effect sizes as the target began to move, both in the pursuit and the saccade tasks (Fig. 3A and B, Extended Data Fig. 5). This increase was the most pronounced in the SNpr (Pursuit – Permutation Welch's ANOVA test: $p<0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}purkinje}$, $p_{SNpr\text{-}local\ cerebellar}<0.001$; Saccades - Permutation Welch's ANOVA test: $p<0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}purkinje}<0.001$, $p_{SNpr\text{-}local\ cerebellar}=0.15$). On the saccade task, the difference between the SNpr and local cerebellar population did not reach significance, but the average effect size was larger in the SNpr. The larger average effect size in the SNpr in comparison to the cerebellar output was surprising, given that the vermis is thought to be functionally closer to the behavior[16,18]. This again violates the expectation derived from a hierarchical organization of reward and movement information across structures.

The coding of direction during the movement epoch was substantially larger than the coding of reward probability in all populations. This was indicated by larger direction effect sizes than reward probability effect sizes (Fig. 3C and Extended Data Fig. 5I-L; Bootstrap t-test: p_{SNpr}, $p_{caudate}$, $p_{purkinje}$, $p_{local\ cerebellar}$<0.001). These results are consistent with previous studies in the basal ganglia[34], cerebellum[11], and frontal eye field[35], implying a general organization of the motor systems in which reward information is mostly represented when it is first available and then only weakly during movement.

Our analysis method allowed us to quantify the extent to which the reward expectation influenced the coding of direction. Neurons in which the reward probability condition substantially modulates the coding of direction should display high values of reward probability x direction interaction effect sizes. These modulations can arise from sharpening or gain modulations of the tuning curve[36]. We did not find such interactions in any of the populations (not shown), suggesting that at least on this task, reward and direction coding influence one another minimally (not shown; Bootstrap t-test: p_{SNpr}=0.18, $p_{caudate}$=0.45, $p_{purkinje}$=0.95, $p_{local\ cerebellar}$=0.31). Thus, the separation in the timing of the coding of reward and movement and the lack of interaction during movement indicate that areas multiplex reward and movement mostly across task epochs rather than in instantaneous activity.

Common control for saccades and pursuit

A substantial fraction of neurons were recorded in both the pursuit and saccades tasks (see Methods). We used ω_p^2 to revisit the question of common control for saccade and pursuit[37] in the vermis[16] and the SNpr[25] and examined this question for the first time in the caudate. We found significant correlations between the direction effect sizes in the pursuit and saccade tasks in all four populations (Fig. 3D and E; Spearman correlation: r_{SNpr}=0.48, $r_{caudate}$=0.59, $r_{purkinje}$=0.65, $r_{local\ cerebellar}$=0.64, p<0.001 in all populations). These correlations indicate that in the vermis and both the input and output structures of the basal ganglia, neurons that are tuned to direction during pursuit also tend to be tuned to direction during saccades. Thus, our results support common control for saccades and pursuit in the subcortical networks.

The coding of reward outcome is comparable in the cerebellum and the basal ganglia

Reward outcome and prediction error signals were found in the basal ganglia[5] and recently in the cerebellum[1,9,38,39]. Here, we used the trials in which reward was delivered or omitted to directly compare the reward outcome signals in the basal ganglia and the cerebellum. We then used the probabilistic design of the tasks to test whether the response to the reward outcome was modulated by the predictability of the reward. Similarly to the cue epoch analysis, we pooled neurons across tasks.

We calculated the ω_p^2 to partition the responses into reward outcome, reward probability, and direction during the outcome epoch (examples in Extended Data Fig. 6). As found during movement, the reward probability effect sizes during the outcome epoch were small in comparison to the effect sizes during the cue epoch, but reached significance in all populations (Fig. 4A; Bootstrap t-test: p<0.001, in all populations). The average direction effect sizes of the population remained in the same order as when aligned to the target motion (Fig. 4B; Permutation Welch's ANOVA test: p<0.001, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}}$

$_{purkinje}$, $p_{SNpr\text{-}local\ cerebellar}$<0.001). Coding of direction was expected since monkeys typically saccade back from the different eccentric positions at the end of the trial (saccade latency: 235±76 ms for Monkey A, 268±74 ms for Monkey G). Thus, finding the largest coding of movement at the outcome epoch in the SNpr supports our conclusion that coding of eye movement direction is the most pronounced in the output of the basal ganglia, contradicting a hierarchical organization of the subcortical network.

We found that the average reward outcome effect size was slightly higher in the cerebellum than in the caudate and the SNpr (Fig. 4C; Permutation Welch's ANOVA test: $p<0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}cerebellum}$, $p_{caudate\text{-}cerebellum}<0.001$, $p_{Purkinje\text{-}cerebellar\ local}=0.56$). This result contrasted sharply with the coding of reward probability during the cue (Fig. 3A) in which the SNpr exhibited the most prominent response. We return to this point in the discussion.

According to the reward prediction error hypothesis, reward predictability should modulate the response to the reward outcome[40–43]. In terms of effect size, prediction error or any outcome modulations based on reward predictability (such as signed or unsigned prediction errors) should be expressed in a positive reward probability x reward outcome interaction effect size. However, we failed to find this interaction in any of the populations (Fig. 4D; Bootstrap t-test: $p_{SNpr}=0.11$, $p_{caudate}=0.09$, $p_{local\ cerebellar}=0.54$, $p_{purkinje}=0.49$). The coding of reward outcome and probability were stronger than the coding of the interaction (Fig. 4E and F; Bootstrap t-test: $p<0.01$ in all populations) indicating that reward prediction error coding and any other interaction between probability and outcome were at best only weakly coded in the recorded populations. This suggests there may be a different functional role for the recorded populations as compared to dopaminergic neurons[2] or climbing fibers[8,10] that code the reward prediction error.

<u>Coding of reward and movement in cerebellar complex spikes</u>

Purkinje cells have the unique property of displaying two distinct types of spikes: simple spikes that are similar to spikes observed in other neurons, which we discussed above, and complex spikes that occur at a much lower frequency (~1Hz) and are hypothesized to serve as teaching signals that affect plasticity[6,7] (see example in Fig. 1F). Recently, complex spikes have been shown to code for reward signals[8–11]. We therefore extended our analysis to complex spikes recorded in the Purkinje cell population.

The complex spike reward probability effect sizes during the cue epoch (Figure 5A; Bootstrap t-test: $p<0.05$) and reward outcome during the outcome epoch (Figure 5B; Bootstrap t-test: $p<0.05$) were significant but small. During the target motion epoch, the effect size of the reward was not significantly different from 0, consistent with previous results[11] (not shown; Bootstrap t-test: $p=0.7$). The coding of direction during the motion epoch on the pursuit (Figure 5C; Bootstrap t-test: $p<0.001$) and saccade (Figure 5D; Bootstrap t-test: $p<0.05$, $n=73$) tasks was also small relative to all other populations. A paired comparison of the effect size of the complex and simple spikes from the same neuron indicated that for all task variables the effect size was substantially larger for the simple spikes (scatter plots in Fig. 5A-D, Bootstrap t-test: $p<0.05$ in all comparisons).

Note, however, that effect size characterizes the signal-to-noise ratio of single neurons. Thus, a large variance relative to the effect will result in a small effect size. We confirmed that the population response (which averages out single trial and single neuron noise) was modulated by the direction of movements (Extended Data Fig. 6). As shown above, the directional response of complex spikes, was not modulated by reward expectation[11]. Thus, although complex spikes code for reward and direction, their low signal-to-noise ratio limits their ability to transform information at the single neuron level. Instead, they may serve as a teaching signal that promotes plasticity[6,7] within the cerebellum to change the simple spike coding of reward or movement.

Controlling for sampling bias

One of the challenges of comparing coding across populations or structures is that the populations themselves may not be homogenous. Thus, some sub-populations may be oversampled and others undersampled. Specifically, in the current study, non-relevant neurons could have decreased the average effect size, whereas some task-relevant neurons could have been missing from our sample. This problem was overcome to some extent by focusing solely on eye-movement areas within the basal ganglia and cerebellum. Thus, we compared subsets of neurons that are very likely to participate in the task and are part of the same system. To further verify that our conclusions were not the result of sampling error we replicated the analysis solely on neurons that showed a time-varying response to the task (see Methods, Extended Data Fig. 8).

To verify that our results were not specific to the vermis, we also extended our analysis to an additional eye movement area within the cerebellum, the flocculus complex (Extended Data Fig. 9). In a previous report, we examined the coding of reward size and movement in the floccular complex and its neighboring areas[11] in a similar task to the pursuit task, in which the color of the cue indicated the future reward size instead of its probability. We reanalysis of the floccular complex data and found that the reward effect sizes in the floccular complex were comparable to the vermis reward probability effect sizes (Extended Data Fig. 9). In addition, although some local floccular neurons had comparable direction effect sizes to neurons in the SNpr, the average effect size was larger in the SNpr than in the floccular complex.

Discussion

Our results extend recent findings on reward information in the cerebellum[1,8–11] and relate them directly to signals found in the basal ganglia[5,31,44]. In comparison to the basal ganglia output structure, reward expectancy (Fig. 2) and sensorimotor (Fig. 3) signals in the cerebellum were small. Reward outcome signals in the cerebellum were comparable to the basal ganglia (Fig. 4), suggesting that the reward consequences of behavior are processed in both areas similarly, although we cannot completely rule out the possibility that some of the activity is related to food consumption movements following reward delivery. The findings also showed that reward outcome signals were not modulated by reward predictability (Fig. 4), suggesting that at reward delivery the recorded populations are more strongly related to the processing of the current reward outcome than the prediction error.

The motor system is organized hierarchically. Neurons closer anatomically to the motor periphery show more resemblance to the final motor command[14,45–48]. In terms of transforming reward information into movement, a hierarchical organization of the motor system suggests that reward should be more strongly represented in the upstream areas and that motor parameters, such as eye movement direction should be more strongly represented in downstream areas. Our results do not support this type of organization across the subcortical networks. In particular, the SNpr, which is anatomically and functionally more remote than the vermis to the motor command[15,16,18] exhibited a stronger representation of both reward expectation and movement (Fig. 2 and 3). Other inconsistencies with a hierarchical organization emerged from the comparable coding of the reward outcome between the basal ganglia and cerebellum (Fig. 4).

Our results also run counter to the assumption of a local hierarchical organization within the cerebellum or the basal ganglia. A local hierarchal structure predicts that reward will be more strongly coded by local or input neurons (such as the caudate or local neurons in the vermis), and direction by output neurons (such as the SNpr or Purkinje cells). However, in both areas, the flow of information was different.

In the basal ganglia, we found amplification of both reward and movement direction signals from the caudate to the SNpr (Fig. 2 and 3). This type of enhancement can be achieved by the convergence of noisy neurons onto a single SNpr neuron to reduce noise. Anatomically, there is a strong convergence from input to output in the basal ganglia that can support this function[49]. Thus, in the basal ganglia, noisy high-dimensional representations are transformed into low-dimensional representations with high signal-to-noise ratios. This extraction of task-relevant information is akin to computational algorithms that compress information, such as principal component analysis. Based mostly on anatomy, it has been hypothesized that the basal ganglia implement these computations[49,50]. Thus, our results provide supporting evidence for this hypothesis from the activity of neurons. We note that the response of the caudate neurons in our sample was delayed relative to the SNpr (Fig. 2A and Fig. 3A and B), implying additional inputs to the SNpr either from within the striatum or from other sources[51]. In the cerebellum, the coding of both reward and direction variables was comparable between output and the local neuron populations. This suggests that although Purkinje cells have the largest convergence factor in the central nervous system, this convergence does not result in an averaging out of task-irrelevant variability.

To put forward and gain insights into network factors that could explain the difference in signal sharpening in the basal ganglia and cerebellum, we studied signal convergence from a group of input neurons to a single output neuron in a model (Fig. 6A, see Methods). Input neurons can be thought of as caudate or local cerebellar neurons that project to the same output SNpr or Purkinje cell. In this model, the trial-by-trial variability in the activity was split into the variance related to the signal and to the noise. The inputs were divided into two subgroups, with opposite representations of the signal (red or blue parts in Fig. 6A). In the context of our task, these subsets correspond to neurons with a higher rate in the P=0.75 condition and neurons with a higher rate in the P=0.25 condition, or to neurons with opposite preferred directions. The input neurons are also influenced by several sources of noise that result in activity that is correlated between neurons but is unrelated to the signal (grey shadings). The neurons converge into an output neuron, which averages its inputs.

We varied two parameters in the model. The first was the heterogeneity of the inputs, which we manipulated by varying the fraction of input neurons with the same polarity of signal

representation that project to the same output neuron (e.g., the fraction of neurons with higher rates in the P=0.75 condition relative to the P=0.25). Intuitively, more heterogeneity in the inputs should decrease the signal-to-noise ratio of the output since opposing signals are averaged out. The second parameter was the degree of noise independence, which we varied by changing the number of independent noise sources. Intuitively, a larger number of independent noise sources should result in a larger signal-to-noise ratio since the noise sources will average out.

Fig. 6B shows the degree of signal sharpening as a function of signal heterogeneity and noise independence. We defined the signal sharpening as the ratio between the signal-to-noise ratio of neurons in the input and output. We used the effect size analysis in the cue epoch to plot the basal ganglia and cerebellum signal sharpening (contour lines in Fig. 6B, see Methods). The two populations were consistent with different regimes of the parameters of the model. For any degree of noise dependence, the cerebellum curve had a larger degree of signal heterogeneity. This pattern of connectivity does not substantially improve the signal-to-noise ratio. However, the diverse inputs have an advantage during learning as they make connections available for rapid learning of new reward associations or adjustment of movements[52]. We have constructed the model to put forward the network logic that could underlie our results. Note that we ignored implementation parameters such as the internal circuitry in each structure or single neuron properties. Thus, the model demonstrates that the sharpening of signal could result from network properties and provides a framework for further research of the basal ganglia and cerebellum.

Methods

We collected neural and behavioral data from one male (monkey A) and one female (monkey G) Macaca Fascicularis monkeys (3.5-5.5 kg). All procedures were approved in advance by the Institutional Animal Care and Use Committees of the Hebrew University of Jerusalem and were in strict compliance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals. We first implanted head holders to restrain the monkeys' heads in the experiments. After the monkeys recovered from surgery, they were trained to sit calmly in a primate chair (Crist Instruments) and consume liquid food rewards (baby food mixed with water and infant formula) from a tube set in front of them. We trained the monkeys to track spots of light that moved across a video monitor placed in front of them.

Visual stimuli were displayed on a monitor (55 and 63 cm from the eye of monkeys A and G). The stimuli appeared on dark background in a dimly lit room. A computer performed all real-time operations and controlled the sequences of target motions. The position of the eye was measured with a high temporal resolution camera (1 kHz, Eyelink 1000 plus- SR research) and collected for further analysis. Monkeys received a reward when tracking the target successfully. We performed two subsequent surgeries to place a round 19 mm diameter recording cylinder over the basal ganglia and the vermis. We began by recording from the caudate and then moved to the SNpr, while continuously recording from the vermis.

Up to 5 quartz-insulated tungsten electrodes (impedance of 1-2 Mohm) were lowered through each cylinder using the Mini-Matrix System (Thomas Recording GmbH). When lowering the electrodes, we searched for neurons that responded during pursuit or saccade eye movements, but sometimes collected data from neurons that did not respond to eye movements. We classified neurons as Purkinje cells by the presence of complex spikes. Table

1 shows the number of neurons recorded from each population. In the analyses of the cue and outcome epochs, we pooled neurons across tasks. If a neuron was recorded in more than one task, we included the task in which the neuron was recorded for more trials. When calculating the complex spike average responses (Extended Data Fig. 7), we reduced the minimal trial number limit to 50, resulting in 15 significantly directionally tuned complex spikes in the saccade task, and 17 in the pursuit task.

	Pursuit task	Saccade task	Both saccade and pursuit tasks
Caudate	86/76	63/108	41/62
SNpr	62/104	38/112	34/52
Local cerebellar neurons	61/42	37/39	30/31
Purkinje cell simple spikes	40/33	29/50	22/31
Purkinje cell complex spikes	56/29	35/38	64/46
Simple and complex spike pairs	36/23	27/26	38/29

Table 1: Number of neurons recorded from each monkey in each task. The Left and right values show the number of neurons for Monkey A and G.

Signals were digitized at a sampling rate of 40 kHz (OmniPlex, Plexon). For the detailed data analysis, we sorted spikes offline (Plexon). For sorting, we used principal component analysis and corrected manually for errors. In some of the Purkinje cells, the complex spikes had distinct low-frequency components. In these cells, we used low-frequency features to identify and sort the complex spikes[53,54]. We paid special attention to the isolation of spikes from single neurons. We visually inspected the waveforms in the principal component space and only included neurons for further analysis when they formed distinct clusters. The sorted spikes were converted into time-stamps with a time resolution of 1 ms and were inspected again visually to check for instability and obvious sorting errors.

We also recorded licking behavior to control for behavioral differences between reward conditions that might confound the results. Licks were recorded using an infra-red beam. Monkey A tended not to extend its tongue, therefore we recorded its lip movements.

Experimental design

Before beginning the neural recordings, we examined the behavior of the monkeys under different reward probability conditions (Extended Data Fig. 1). Our selection of probabilities of 0.25 and 0.75 for recordings has several advantages. First, the large difference between probabilities prompts a large difference in reward expectation before delivery. On another task in which the monkeys were required to choose between two targets, the monkeys consistently chose the 0.75 probability target (Fig. 1C) indicating that they associated the color of the target with the probability of reward. Second, a probabilistic reward made it possible to test for differences in neural activity between reward delivery and omission (reward outcome) and manipulate reward expectation at delivery to test whether the neural activity was consistent with the reward prediction error hypothesis[40]. Third, the effect of the reward probability on pursuit and saccades was small in comparison to a condition in which

428 the monkey never received a reward (P=0, Extended Data Fig. 1), and in comparison, to the
429 variability in the behavior within conditions. This similarity in the behavior served to examine
430 the response to reward information separately from movement preparation and execution.

431 **Pursuit Task:**

432 Each trial started with a bright white target that appeared in the center of the screen
433 (Figure 1A). After 500 ms of presentation, in which the monkey was required to acquire
434 fixation, a colored target replaced the fixation target. The color of the target signaled the
435 probability of receiving a reward upon successful tracking of the target. For monkey A we used
436 yellow to indicate a 0.75 probability for reward and green to indicate 0.25. For monkey G we
437 reversed the associations. After a variable delay of 800-1200 ms, the targets stepped in one
438 of eight directions (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°) and then continuously moved in
439 the opposite direction (step-ramp)[19]. For both monkeys, we used a target motion of 20°/s and
440 a step to a position 4° from the center of the screen. The target moved for 750 ms for monkey
441 A and 650 ms for monkey G and then stopped and stayed still for an additional 500-700 ms. If
442 the monkey's gaze was within a 3-5°x3-5° window around the target, the monkey received a
443 reward according to the probability specified by the color.

444 **Saccade Task:**

445 The structure of the saccade task was identical to the pursuit task, except for the
446 target motion epoch. In the saccade task, following the random delay, the central colored
447 target disappeared and immediately reappeared in one of 8 eccentric locations (0°, 45°, 90°,
448 135°, 180°, 225°, 270°, 315°) 10° from the center of the screen. If the monkey's gaze was within
449 a 5°x5° window around the target, the monkey received a reward according to the probability
450 specified by the color.

451 **Choice Task:**

452 We used this task to verify that the monkeys associated the color of the target with
453 the reward probability. The monkeys were required to choose one of two targets with
454 different reward probability values presented on the screen. The P=0.25 and P=0.75 color
455 associations were identical to the pursuit and saccade tasks. The structure of the task was
456 similar to previous reports[55].

457 **Flocculus reward size task:**

458 We used previously analyzed data recorded from the floccular complex and adjacent
459 areas, while the monkeys performed a smooth pursuit task in which we manipulated reward
460 size. The full details can be found in previous reports[11,56]. Briefly, the temporal and target
461 motion properties of the task were the same as the reward probability pursuit task described
462 above. However, in this task, the color of the target signaled the size of the reward the monkey
463 would receive if it tracked the target. One color was associated with a large reward (~0.2ml)
464 and the other with a small reward (~0.05 ml).

465 *Data analysis*

466 All analyses were performed using MATLAB (Mathworks). We only included cells that

were recorded for at least 55 successful trials in either the saccade or pursuit task, with a median of 141 trials.

Eye movement analysis:

To calculate the average of the smooth pursuit initiation, we first removed the saccades and treated them as missing data. We then averaged the traces in the target movement direction. Finally, we smoothed the traces using a Gaussian filter with a standard deviation of 15 ms. We used eye velocity and acceleration thresholds to detect saccades automatically and then verified the automatic detection by visual inspection of the traces. The velocity and acceleration signals were obtained by digitally differentiating the position signal after we smoothed it with a Gaussian filter with a standard deviation of 5 ms. Saccades were defined as an eye acceleration exceeding 1000°/s^2, an eye velocity crossing 15°/s during fixation, or an eye velocity crossing 50°/s while the target moved on the pursuit task.

Peristimulus time histogram:

To study the average time-varying properties of the response, we calculated the peristimulus time histogram (PSTH) at a 1 ms resolution. We then smoothed the PSTH with a 15 ms standard deviation Gaussian window, removing at least 100 ms before and after the displayed time interval to avoid edge effects.

Tuning curves:

To calculate the tuning curves (Extended Data Fig. 7), we averaged the responses in the first 100-300 ms of the movement. We calculated the preferred direction of the neuron as the direction that was the closest to the vector average of the responses across directions (direction of the center of mass). We used the preferred direction to calculate the population tuning curve by aligning all the responses to the preferred direction.

Effect size in time:

To quantify a neuron's coding of a variable through time, we calculated the ω_p^2 effect size in each 50 ms bin separately. In each bin and each trial, we counted the number of spikes, resulting in a distribution of the number of spikes emitted by the neuron in each condition. ω_p^2 is a common effect size measure used in ANOVA designs that is often preferred over other effect size measures since it is unbiased[28,57,58]. The calculation is as follows:

$$\omega_p^2 = \frac{SS_{effect} - \frac{df_{effect}}{df_{error}} \cdot SS_{error}}{SS_{effect} + \frac{\left(N - df_{effect}\right)}{df_{error}} \cdot SS_{error}}$$

where SS_{effect} is the ANOVA sum of squares for the effect of a specific variable, SS_{error} is the sum of squares of the errors after accounting for all experimental variables, df_{effect} and df_{error} are the degrees of freedom for the variable and the error respectively and N is the number of observations (number of trials). During the cue epoch, the reward probability was the only variable in the ANOVA design, during the target motion the variables were the reward probability and direction, and during the outcome epoch, the variables were rewarded probability, direction, and the reward outcome. We also included all possible interactions. We

used type II sum of squares to account for unequal sizes of groups (for example, a smaller number of trials in the reward-delivered as compared to the reward-omitted within the 0.25 condition)[59]. We used the partial effect size since it enables a better comparison between neurons that respond or do not respond to other experimental variables. Using the full effect size did not alter any of our conclusions.

Effect size in the entire epoch:

To calculate the coding of a variable throughout the entire epoch, we calculate ω_p^2 in models that included time as an additional variable. Again, we calculated the number of spikes in 50 ms bins, 0 to 800 ms after an event during the trial (cue appearance, target motion, or outcome). Changing the bin size yielded highly correlated effect sizes (Extended Data Fig. 2D-I). We fitted an ANOVA model that included the same variables as listed above for each epoch, with the addition of time (the specific bin the sample came from). Again, we included all possible interactions. The formula that we used was similar:

$$\omega_p^2 = \frac{SS_{effect} + SS_{effect\ x\ time} - \frac{df_{effect} + df_{effect\ x\ time}}{df_{error}} \cdot SS_{error}}{SS_{effect\ x\ time} + \frac{\left(N - df_{effect} - df_{effect\ x\ time}\right)}{df_{error}} \cdot SS_{error}}$$

where $SS_{effect\ x\ time}$ is the ANOVA type II sum of squares for the interaction of a specific variable with time, and $df_{effect\ x\ time}$ are the corresponding degrees of freedom. We included the interaction term since it quantifies the time-varying coding of the variable. In this case, N is equal to the number of trials x the number of time bins.

We verified that ω_p^2 was not determined by the firing rate of neurons (Extended Data Fig. 3). Although Purkinje cells had the high firing rates, their effect sizes during the cue and target motion epochs were relatively small. Additionally, in some cases rate and effect size were correlated, this correlation was small or non-existent in others.

Time-varying response:

We repeated our analysis on subsets of neurons within each population that significantly changed their average firing rate throughout the epoch (Extended Data Fig. 8). To find these neurons we calculated the significance of the time variable, using the same ANOVA design described above in the subsection Effect size in the entire epoch.

	Cue	Movement: pursuit	Movement: saccade	Outcome
Caudate	92 (0.39)	71 (0.44)	105 (0.61)	171 (0.73)
SNpr	162 (0.7)	146 (0.87)	123 (0.82)	213 (0.92)
Local cerebellar neurons	78 (0.58)	70 (0.67)	66 (0.75)	118 (0.88)
Purkinje cell simple spikes	49 (0.52)	49 (0.67)	46 (0.67)	89 (0.94)

Table 2: Number and fraction (in parenthesis) of neurons that were included in the analysis presented in Extended Data Fig. 8.

Comparison to a classifier:

We compared the reward probability effect size to the 10-fold cross-validated accuracy of a classifier for the reward condition (Extended Data Fig. 2). To train the classifier, we calculated the PSTH of the training set for each reward condition. To predict the label of each trial in the test set, we calculated its single trial smoothed PSTH. We compared the PSTH to the PSTHs calculated over the training set and assigned it the label of the PSTH most similar to it (minimal L_2 norm).

Statistical tests and corrections for multiple comparisons:

The distributions of effect sizes tended to be non-normal and the variances were different between the populations. We therefore avoided using standard ANOVA tests. Instead, we used permutation and bootstrap methods to evaluate the distributions of statistics that correct for unequal variances. When permuting we recalculate the statistics 10,000 times, as we observed that the p-values were stable at this number of repetitions.

Permutation Welch's ANOVA test: We permuted the population labels of the effect sizes and calculated the Welch's ANOVA F statistic for each permutation. The p-value was defined as the probability for the F statistic to be larger than the F statistic for the real sample.

Permutation Welch's t-test: We permuted the population labels of the effect sizes and calculated the Welch's t-test statistic for each permutation. The p-value was defined as the probability for the t statistic to be more extreme than the t statistic for the real sample.

Bootstrap t-test: We used this to test the null hypothesis that effect size values, or the paired differences in effect size values, are significantly different than 0. We resampled (with repetition) from our original sample and subtracted the mean for each sample to simulate the null hypothesis. The p-value was defined as the probability for the t statistic to be farther than 0 from the t statistic of the real sample.

Convergence model:

We modeled the convergence of inputs using a model for convergence from input neurons to an output neuron. The input neurons represent the subgroup of neurons in an input or local structure that project to a specific neuron in the output structure. The activity of each input neuron is described by

$$x_i = \pm\sqrt{\alpha}\mu + \sqrt{1-\alpha}\varepsilon_i$$

where i is the index of the neuron out of a population of N and $0 \leq \alpha \leq 1$ is a parameter determining the signal-to-noise ratio (SNR) in the input neurons. $\mu \in \{-1,1\}$ represents the signal and is distributed across trials with p=0.5. The input neurons are divided into two groups, with opposing representations of μ (hence the $\pm$ in the equations). The sizes of the groups are N_1 and $N_2 = N - N_1$, and in each trial all neurons within the same group will have the same representation of the signal, which will be opposing to the neurons in the other group. ε_i is a noise term that is shared within K groups of neurons. For $K = 1$, the noise is shared across the entire population. For $K = N$, the noise of any pair of neurons is independent. For each of the groups $\varepsilon \sim Normal(0,1)$. We denote the shared error term for group k as $\tilde{\varepsilon}_k$.

The signal-to-noise ratio (SNR) of the input neurons is defined as the ratio between the

variance from μ and the variance from ε_i.

$$SNR(x_i) = \frac{\alpha}{1-\alpha}$$

The activity of the output neuron is described by

$$\begin{aligned} y_i &= \frac{1}{N}\sum_{i=1}^{N} x_i \\ &= \frac{\sqrt{\alpha}}{N}\left(\sum_{i=1}^{N_1}\mu - \sum_{i=1}^{N_2}\mu\right) + \frac{\sqrt{1-\alpha}}{N}\left(\sum_{i=1}^{N}\varepsilon_i\right) \end{aligned}$$

We switch to summing ε by noise group instead of neuron and assumed that the groups are of equal size $\frac{N}{K}$, thus:

$$\begin{aligned} y_i &= \frac{\sqrt{\alpha}(N_1 - N_2)}{N}\mu + \frac{\sqrt{1-\alpha}}{N}\left(\sum_{k=1}^{K}\frac{N}{K}\tilde{\varepsilon}_k\right) \\ &= \frac{\sqrt{\alpha}(N_1 - N_2)}{N}\mu + \frac{\sqrt{1-\alpha}}{K}\left(\sum_{k=1}^{K}\tilde{\varepsilon}_k\right) \end{aligned}$$

The variance of y_i is:

$$\begin{aligned} Var(y_i) &= \alpha\left(\frac{N_1 - N_2}{N}\right)^2 + \frac{1-\alpha}{K^2}\sum_{k=1}^{K} Var(\tilde{\varepsilon}_k) \\ &= \alpha\left(\frac{N_1 - N_2}{N}\right)^2 + (1-\alpha)\frac{1}{K} \end{aligned}$$

The SNR of the output neuron is therefore

$$SNR(y_i) = \frac{\left(\frac{N_1 - N_2}{N}\right)^2}{\frac{1}{K}} \cdot \frac{\alpha}{1-\alpha} = \frac{K(N_1 - N_2)^2}{N^2} \cdot \frac{\alpha}{1-\alpha}$$

Thus, the sharpening of the signal following the convergence is:

$$\frac{SNR(y_i)}{SNR(x_i)} = \frac{K(N_1 - N_2)^2}{N^2}$$

Fig. 6B shows $\frac{SNR(y_i)}{SNR(x_i)}$ as function of $\frac{N_1}{N}$ (horizontal) and $\frac{K}{N}$ (vertical).

To plot the basal ganglia and cerebellum on top of the model, we converted ω_p^2 to SNR. ω_p^2 is an estimator for

$$\omega_p^2 \approx \frac{Var(signal)}{Var(signal) + Var(noise)} = \frac{\frac{Var(signal)}{Var(noise)}}{\frac{Var(signal)}{Var(noise)} + 1}$$

Therefore,

$$\omega_p^2 \approx \frac{SNR}{SNR + 1}$$

$$SNR \approx \frac{\omega_p^2}{1 - \omega_p^2}$$

Acknowledgments: This work is dedicated to the memory of Mrs. Lily Safra, a great supporter of brain research. We thank Lottem E. and Fox L. for commenting on early versions of the manuscript. This project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 755745) and the Israeli Scientific Foundation (380/17).

References

1. Wagner, M. J., Kim, T. H., Savall, J., Schnitzer, M. J. & Luo, L. Cerebellar granule cells encode the expectation of reward. *Nature* **544**, 96–100 (2017).

2. Schultz, W. Predictive reward signal of dopamine neurons. *Journal of Neurophysiology* vol. 80 1–27 Preprint at (1998).

3. Lixenberg, A., Yarkoni, M., Botschko, Y. & Joshua, M. Encoding of eye movements explains reward-related activity in cerebellar simple spikes. *J Neurophysiol* **123**, 786–799 (2020).

4. Doya, K. Complementary roles of basal ganglia and cerebellum in learning and motor control. *Curr Opin Neurobiol* **10**, 732–739 (2000).

5. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science (1979)* **275**, 1593–1599 (1997).

6. Albus, J. S. A Theory of Cerebellar Function. *Math Biosci* **10**, 25–61 (1971).

7. Marr, D. A Theory of Cerebellar Cortex. *J Physiol* 437–470 (1969).

8. Heffley, W. *et al.* Coordinated cerebellar climbing fiber activity signals learned sensorimotor predictions. *Nat Neurosci* **21**, 1431–1441 (2018).

9. Heffley, W. & Hull, C. Classical conditioning drives learned reward prediction signals in climbing fibers across the lateral cerebellum. *Elife* **8**, (2019).

10. Kostadinov, D., Beau, M., Pozo, M. B. & Häusser, M. Predictive and reactive reward signals conveyed by climbing fiber inputs to cerebellar Purkinje cells. *Nat Neurosci* **22**, 950–962 (2019).

11. Larry, N., Yarkoni, M., Lixenberg, A. & Joshua, M. Cerebellar climbing fibers encode expected reward size. *Elife* **8**, 1–16 (2019).

12. Carta, I., Chen, C. H., Schott, A. L., Dorizan, S. & Khodakhah, K. Cerebellar modulation of the reward circuitry and social behavior. *Science (1979)* **363**, eaav0581 (2019).

13. Kameda, M., Ohmae, S. & Tanaka, M. Entrained neuronal activity to periodic visual stimuli in the primate striatum compared with the cerebellum. *Elife* **8**, (2019).

14. Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. v. Cortical activity in the null space: permitting preparation without movement. *Nature Neuroscience 2014 17:3* **17**, 440–448 (2014).

15. Krauzlis, R. J. Recasting the Smooth Pursuit Eye Movement System. *Journal of Neurophysiology* vol. 91 591–603 Preprint at https://doi.org/10.1152/jn.00801.2003 (2004).

16. Krauzlis, R. J. & Miles, F. A. Role of the oculomotor vermis in generating pursuit and saccades: Effects of microstimulation. *J Neurophysiol* **80**, 2046–2062 (1998).

17. Lisberger, S. G. Neural basis for motor learning in the vestibuloocular reflex of primates. III. Computational and behavioral analysis of the sites of learning. *https://doi.org/10.1152/jn.1994.72.2.974* **72**, 974–998 (1994).

18. Basso, M. A. & Liu, P. Context-dependent effects of substantia nigra stimulation on eye movements. *J Neurophysiol* **97**, 4129–4142 (2007).

19. Rashbass, C. & Westheimer, G. Independence of conjugate and disjunctive eye movements. *J Physiol* **159**, 361–364 (1961).

20. Hikosaka, O. & Wurtz, R. H. Visual and oculomotor functions of monkey substantia nigra pars reticulata. IV. Relation of substantia nigra to superior colliculus. *J Neurophysiol* **49**, 1285–1301 (1983).

21. Hikosaka, O., Sakamoto, M. & Usui, S. Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. *J Neurophysiol* **61**, 780–798 (1989).

22. Noda, H. & Fujikado, T. Topography of the oculomotor area of the cerebellar vermis in macaques as determined by microstimulation. *https://doi.org/10.1152/jn.1987.58.2.359* **58**, 359–378 (1987).

23. Raghavan, R. T. & Lisberger, S. G. Responses of Purkinje cells in the oculomotor vermis of monkeys during smooth pursuit eye movements and saccades: Comparison with floccular complex. *J Neurophysiol* **118**, 986–1001 (2017).

24. Suzuki, D. A. & Keller, E. L. The role of the posterior vermis of monkey cerebellum in smooth-pursuit eye movement control. II. Target velocity-related Purkinje cell activity. *J Neurophysiol* **59**, 19–40 (1988).

25. Basso, M. A., Pokorny, J. J. & Liu, P. Activity of substantia nigra pars reticulata neurons during smooth pursuit eye movements in monkeys. *European Journal of Neuroscience* **22**, 448–464 (2005).

26. Eccles, J. C., Llinás, R. & Sasaki, K. The excitatory synaptic action of climbing fibres on the Purkinje cells of the cerebellum. *J Physiol* **182**, 268–296 (1966).

27. Stone, L. S. & Lisberger, S. G. Visual responses of Purkinje cells in the cerebellar flocculus during smooth-pursuit eye movements in monkeys. II. Complex spikes. *J Neurophysiol* **63**, 1262–1275 (1990).

28. Olejnik, S. & Algina, J. Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemp Educ Psychol* **25**, 241–286 (2000).

29. Shin, S. Y. & Sommer, M. A. Activity of neurons in monkey globus pallidus during oculomotor behavior compared with that in substantia nigra pars reticulata. *J Neurophysiol* **103**, 1874–1887 (2010).

30. Joshua, M., Adler, A., Rosin, B., Vaadia, E. & Bergman, H. Encoding of probabilistic rewarding and aversive events by pallidal and nigral neurons. *J Neurophysiol* **101**, 758–772 (2009).

31. Nakamura, K. & Hikosaka, O. Role of dopamine in the primate caudate nucleus in reward modulation of saccades. *Journal of Neuroscience* **26**, 5360–5369 (2006).

32. Cromwell, H. C., Hassani, O. K. & Schultz, W. Relative reward processing in primate striatum. *Exp Brain Res* **162**, 520–525 (2005).

33. Churchland, M. M. *et al.* Neural population dynamics during reaching. (2012) doi:10.1038/nature11129.

34. Arkadir, D., Morris, G., Vaadia, E. & Bergman, H. Behavioral/Systems/Cognitive Independent Coding of Movement Direction and Reward Prediction by Single Pallidal Neurons. (2004) doi:10.1523/JNEUROSCI.2583-04.2004.

35. Raghavan, R. T. & Joshua, M. Dissecting patterns of preparatory activity in the frontal eye fields during pursuit target selection. *J Neurophysiol* jn.00317.2017 (2017).

36. Martinez-Trujillo, J. C. & Treue, S. Feature-Based Attention Increases the Selectivity of Population Responses in Primate Visual Cortex. *Current Biology* **14**, 744–751 (2004).

37. Krauzlis, R. J. Activity of rostral superior colliculus neurons during passive and active viewing of motion. *J Neurophysiol* **92**, 949–958 (2004).

38. Kostadinov, D. & Hausser, M. Reward signals in the cerebellum: origins, targets, and functional implications. *Neuron* (2022).

39. Sendhilnathan, N. *et al.* Neural Correlates of Reinforcement Learning in Mid-lateral Cerebellum. *Neuron* **106**, 1–11 (2020).

40. Fiorillo, C. D., Tobler, P. N. & Schultz, W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science (1979)* **299**, 1898–1902 (2003).

41. Morris, G., Arkadir, D., Nevet, A., Vaadia, E. & Bergman, H. Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* **43**, 133–143 (2004).

42. Schultz, W. Predictive reward signal of dopamine neurons. *J Neurophysiol* **80**, 1–27 (1998).

43. Sutton, R. S. & Barto, A. G. Reinforcement Learning: An Introduction. *IEEE Trans Neural Netw* (1998) doi:10.1109/tnn.1998.712192.

44. Sato, M. & Hikosaka, O. Role of Primate Substantia Nigra Pars Reticulata in Reward-Oriented Saccadic Eye Movement. *Journal of Neuroscience* **22**, 2363–2373 (2002).

45. Joshua, M., Medina, J. F. & Lisberger, S. G. Diversity of Neural Responses in the Brainstem during Smooth Pursuit Eye Movements Constrains the Circuit Mechanisms of Neural Integration. *Journal of Neuroscience* **33**, 6633–6647 (2013).

46. Sparks, D. L. The brainstem control of saccadic eye movements. *Nature Reviews Neuroscience 2002 3:12* **3**, 952–964 (2002).

47. Robinson, D. A. The use of control systems analysis in the neurophysiology of eye movements. *An~ Rev. NeuroscL* **4**, 463–503 (1981).

48. Shalit, U., Zinger, N., Joshua, M. & Prut, Y. Descending Systems Translate Transient Cortical Commands into a Sustained Muscle Activation Signal. *Cerebral Cortex* **22**, 1904–1914 (2012).

49. Bar-Gad, I., Morris, G. & Bergman, H. Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Prog Neurobiol* **71**, 439–473 (2003).

50. Bar-Gad, I., Goldberg, J. A., Bergman, H., Havazelet-Heimer, G. & Ruppin, E. Reinforcement-Driven Dimensionality Reduction- A Model For Information Processing In The Basal Ganglia. *J Basic Clin Physiol Pharmacol* **11**, 305–320 (2000).

51. Nambu, A. *et al.* Excitatory conical inputs to pallidal neurons via the subthalamic nucleus in the monkey. *J Neurophysiol* **84**, 289–300 (2000).

52. Jörntell, H. & Ekerot, C. F. Receptive Field Plasticity Profoundly Alters the Cutaneous Parallel Fiber Synaptic Input to Cerebellar Interneurons In Vivo. *Journal of Neuroscience* **23**, 9620–9631 (2003).

53. Sedaghat-Nejad, E. *et al.* P-sort: an open-source software for cerebellar neurophysiology 3 4. doi:10.1101/2021.03.16.435644.

54. Zur, G. & Joshua, M. Using extracellular low frequency signals to improve the spike sorting of cerebellar complex spikes. *J Neurosci Methods* **328**, 108423 (2019).

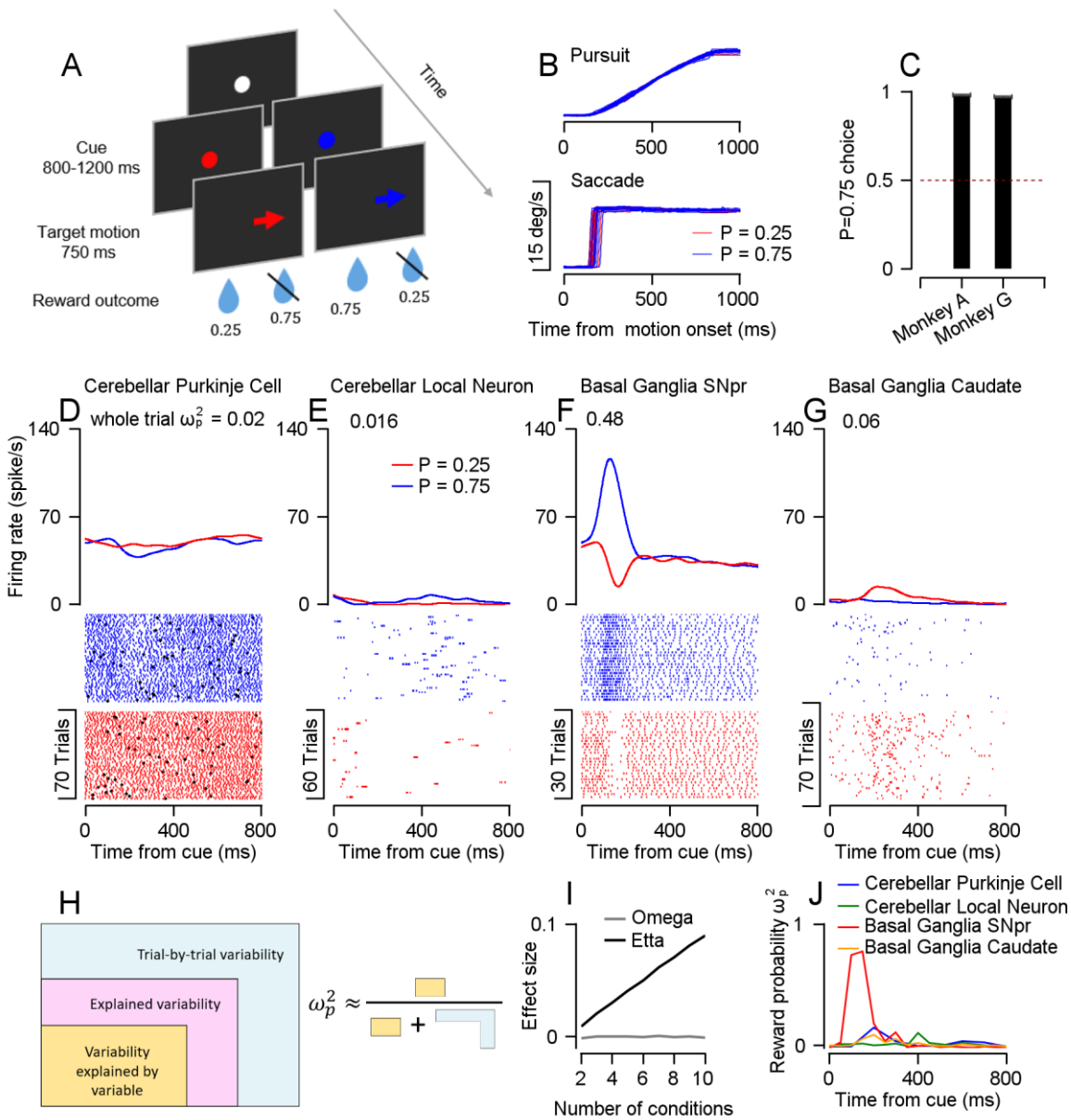
55. Joshua, M. & Lisberger, S. G. Reward Action in the Initiation of Smooth Pursuit Eye Movements. *Journal of Neuroscience* **32**, 2856–2867 (2012).

56. Lixenberg, A. & Joshua, M. Encoding of reward and decoding movement from the frontal eye field during smooth pursuit eye movements. *Journal of Neuroscience* **38**, 10515–10524 (2018).

57. Olejnik, S. & Algina, J. Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. *Psychol Methods* **8**, 434–447 (2003).

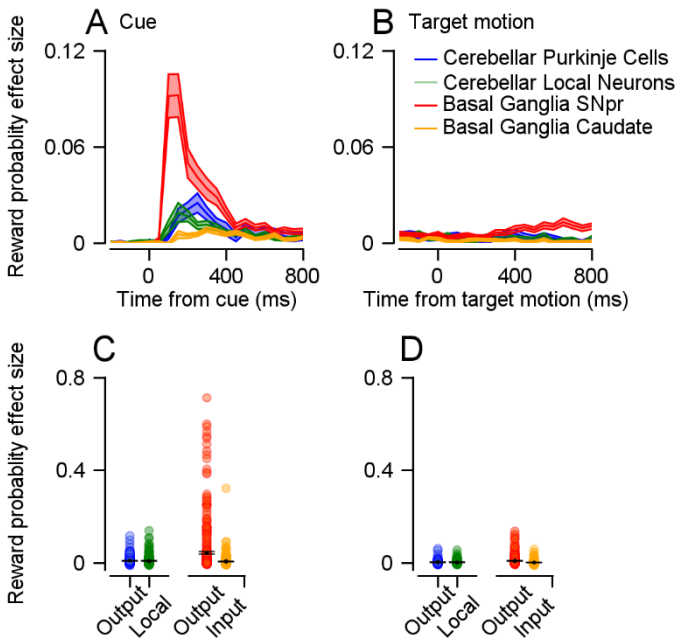
58. Okada, K. Is Omega Squared Less Biased? a Comparison of Three Major Effect Size Indices in One-Way Anova. *Behaviormetrika* **40**, 129–147 (2013).

59. Langsrud, Ø. ANOVA for unbalanced data: Use Type II instead of Type III sums of squares. *Stat Comput* **13**, 163–167 (2003).

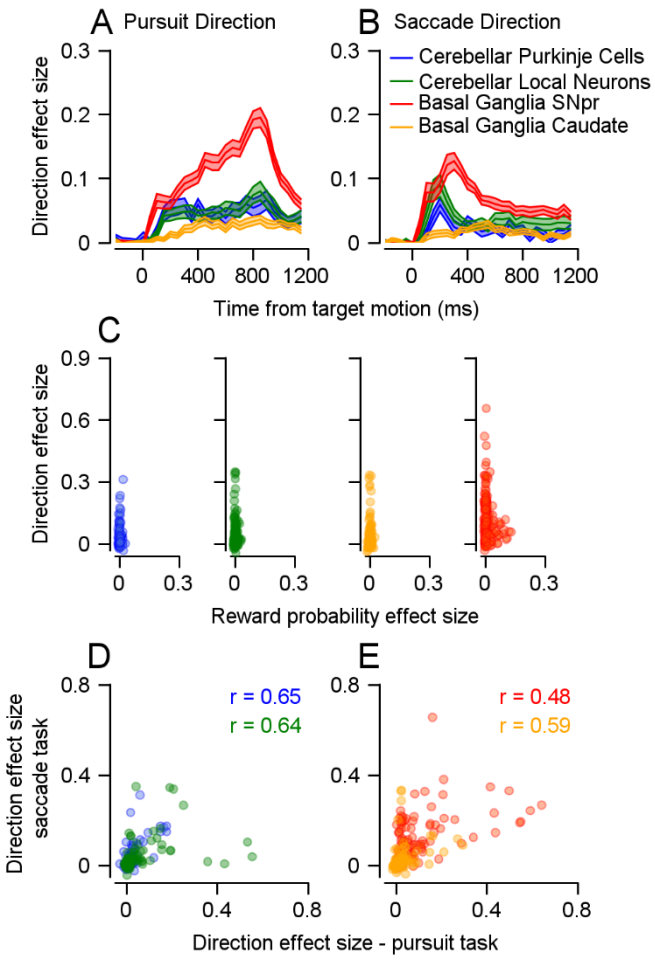


<u>Fig. 1:</u> Effect size analysis to quantify the coding of reward probability on a trial-by-trial basis. A, Schematics of the pursuit and saccade eye movement tasks. The spots represent the target and the arrows show the direction of continuous motion (pursuit) or instantaneous jump of the target (saccade). The numbers represent the probability of receiving a reward. **B,** Ten example traces of eye position from a single pursuit (top) and saccade (bottom) session aligned to target motion onset. **C**, The fraction of trials in which the monkey selected the P=0.75 target on an additional choice task averaged across sessions. The dashed line represents the chance level. All error bars represent the standard error of the mean. **D**-**G**, PSTHs (top) and rasters (bottom) aligned to the onset of the cue for the P=0.25 (red) and P=0.75 (blue) conditions in a cerebellar Purkinje cell (**D**), cerebellar local neuron (**E**), SNpr neuron (**F**), and caudate neuron (**G**). The black diamonds in **F** represent complex spikes recorded in the same Purkinje cell. **H**, An illustration of the calculation of the ω_p^2 effect size. The parts represent the partition of the trial-by-trial variability into the variance explained by a specific variable (yellow), the variance explained by other variables (pink), and the variance unexplained by any of the task variables (blue). **I**, Average ω_p^2 (grey) and partial η^2 (black) effect sizes as a function of the number

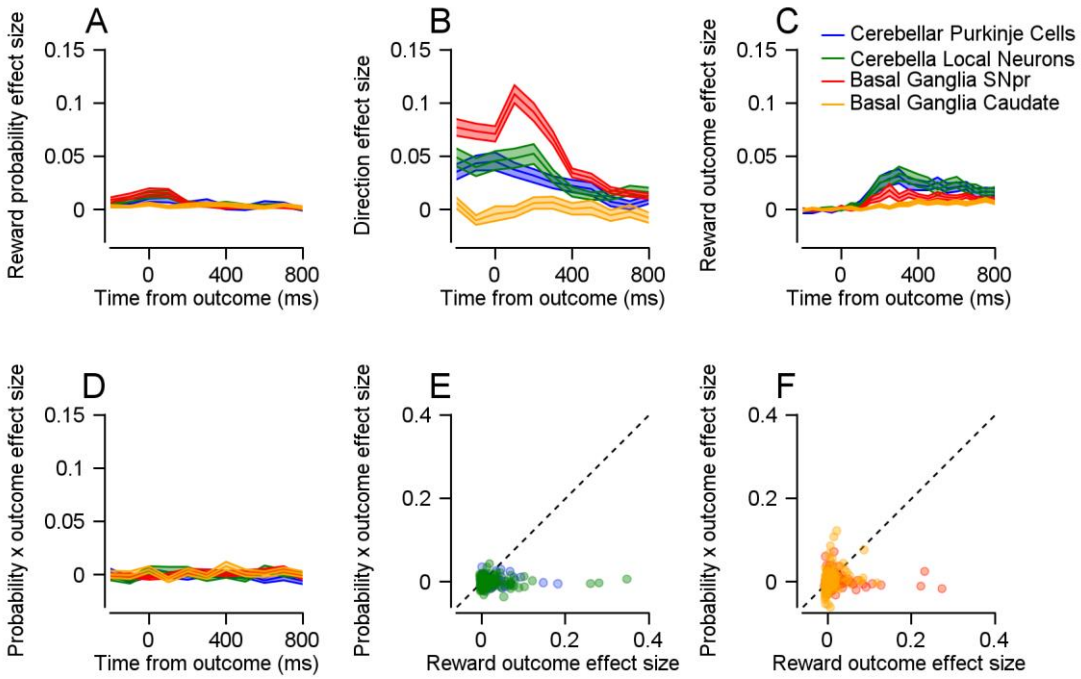
of conditions in a simulation in which the neural responses were independent of the condition. **J**, Reward probability ω_p^2 effect sizes of neurons in **D**-**G** for a model with reward as the sole variable, calculated in bins of 50 ms. Entire trial effect sizes that represent the total coding of reward probability throughout the cue epoch are shown in **D**-**G**.



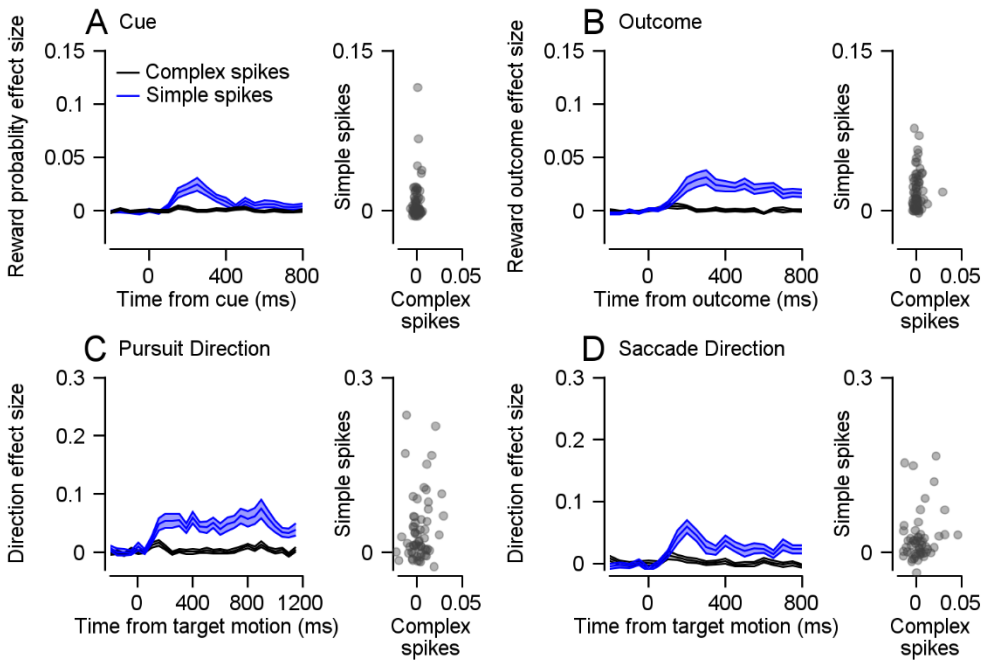
<u>Fig. 2:</u> Coding of reward probability during the cue and movement. A and **B**, Average reward probability effect size in the cue (**A**) and target motion (**B**) epochs. **C** and **D,** Each dot represents a single neuron's reward probability effect size in the cue epoch (**C**) and target motion (**D**) epochs. Bars show the average and standard error of the mean.



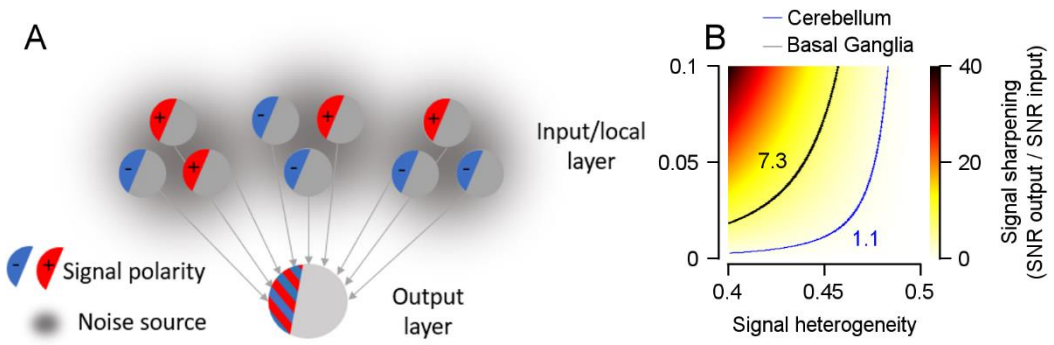
<u>Fig. 3:</u> Coding of eye movement direction during pursuit and saccades. A and **B,** Direction effect size along the target motion epoch in the pursuit task (**A**) and the saccade task (**B**). **C**, Each dot represents a single neuron's reward probability (horizontal) and direction (vertical) effect sizes in the target motion epoch. **D** and **E**, Each dot represents a single neuron's direction effect size in the pursuit (horizontal) and saccade (vertical) tasks. The r values represent the Spearman correlation. **D** shows the basal ganglia populations and **E** the vermis.



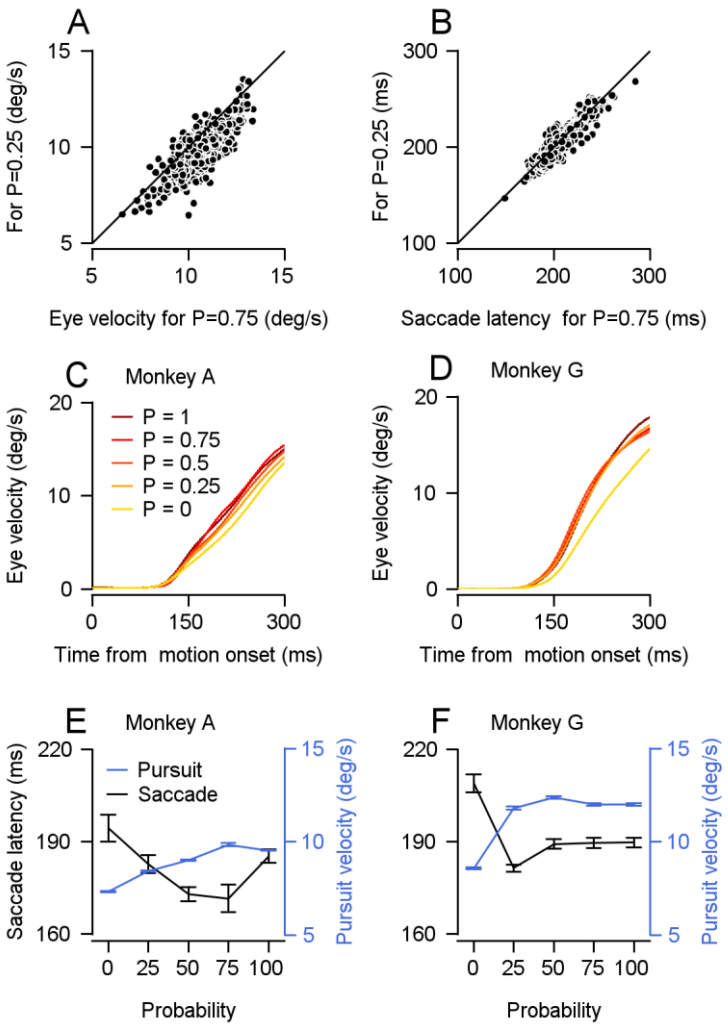
<u>Fig. 4:</u> Coding of reward outcome at the end of the trial. A-D, Reward probability (**A**), direction (**B**), reward outcome (**C**) and reward probability x reward outcome interaction (**D**) effect sizes in the outcome epoch. **E** and **F**, Each dot represents a single neuron's reward outcome effect size (horizontal) and reward probability x reward outcome interaction effect size (vertical). **E** shows the vermis populations and **F** the basal ganglia populations.



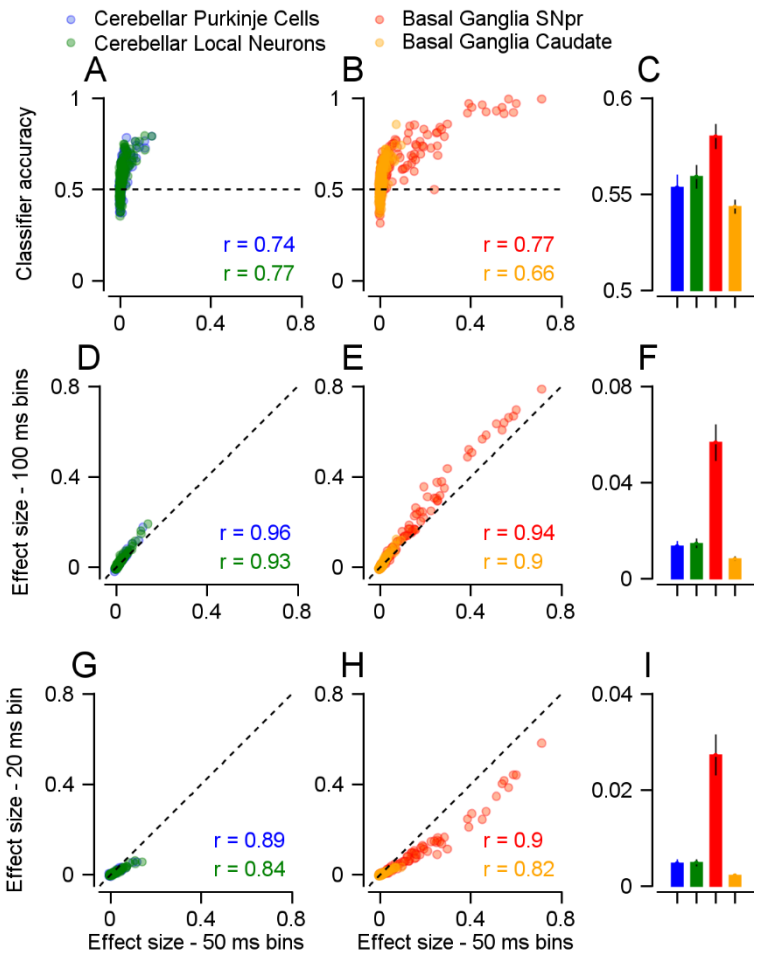
<u>Fig. 5:</u> Effect sizes of complex spikes in the cerebellar vermis. In all figures, black traces represent complex spikes recorded from Purkinje cells in the vermis. The blue traces show the simple spike effect sizes displayed previously. In the scatter plots, each dot represents a single Purkinje cell's complex (horizontal) and simple (vertical) spike effect sizes. **A**, Reward probability effect size in the cue epoch. **B**, Reward outcome effect size in the outcome epoch. **C** and **D**, Direction effect size in the target motion epoch in the pursuit (**C**) and saccade (**D**) tasks.



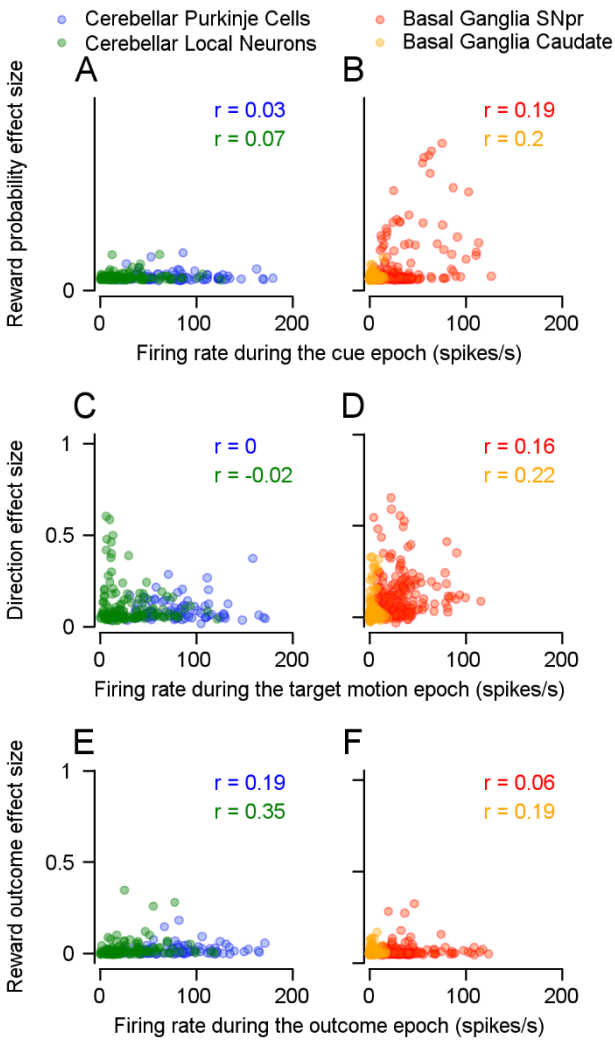
<u>Fig. 6:</u> Convergence model for the basal ganglia and the cerebellum. A, Schematics of the convergence model. Small circles represent input or local neurons. The grey part of the circle represents the variance related to noise and the blue and red parts represent the variance related to signal. Each neuron has a specific signal polarity (plus or minus) and belongs to a set of neurons with correlated noise, indicated by the gray shading in the background. The neurons converge to an output neuron, whose variance can also be portioned into noise and signal. **B**, The color represents the signal sharpening as a function of the fraction of neurons with the same signal polarity (horizontal), and the number of noise groups (vertical). The blue and black curves represent the contour lines of the signal sharpening in the basal ganglia (7.3) and cerebellum (1.1) during the cue epoch.



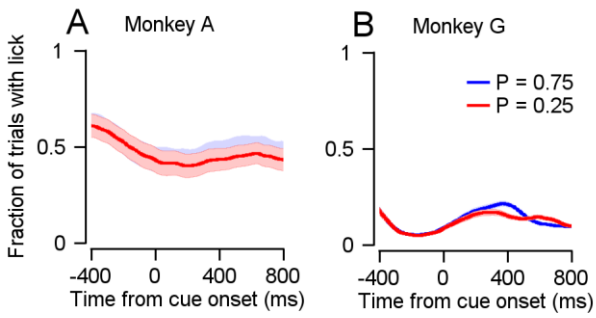
<u>Extended Data Fig. 1:</u> Pursuit velocity and saccade latency are relatively similar in the P=0.25 and P=0.75 reward probability conditions. A, Each dot represents the average speed for an individual session 200 to 250 ms after target motion onset for the high (horizontal) and low (vertical) reward probability (Signed-rank: $p<0.001$, $n=633$). **B**, Each dot represents the average saccade latency for an individual session (Signed-rank: $p<0.001$, $n=565$). **C** and **D**, Traces of average eye velocity in the first 300 ms after target motion onset in sessions with multiple probabilities on the pursuit task for monkeys A (**C**) and G (**D**). **E** and **F**, Black traces correspond to the vertical axis on the left and show the average saccade latencies in sessions with multiple probabilities on the saccade task. Blue traces correspond to the vertical axis on the right and show the average pursuit velocity 200 to 250 ms after the target motion in sessions with multiple probabilities in the pursuit task. **E** and **F** show data from monkeys A and G.



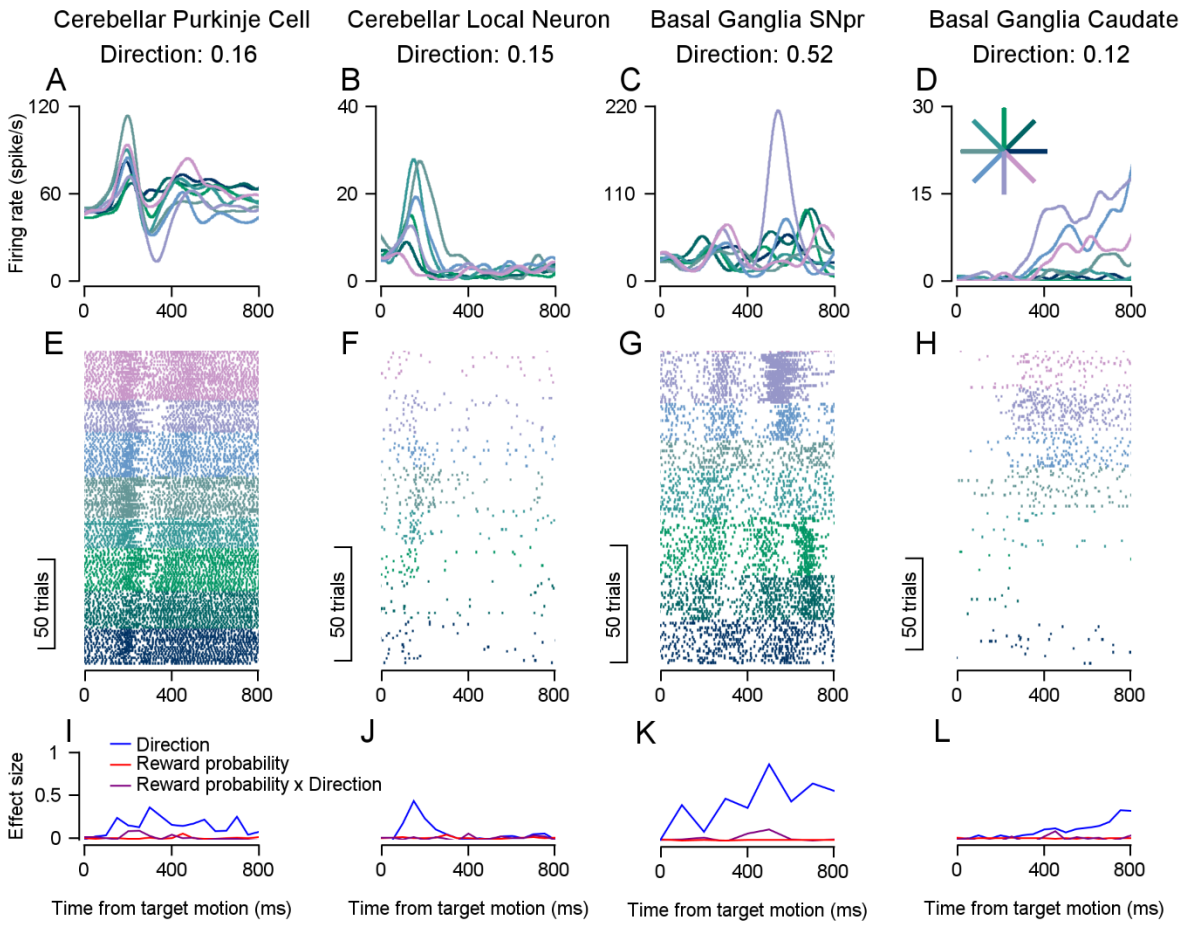
<u>Extended Data Fig. 2:</u> ω_p^2 effect sizes are correlated with the accuracy of the classifier and are correlated across different bin sizes. Each dot represents a single neuron. In all plots, the horizontal position of the dots represents the ω_p^2 reward probability effect size calculated in 50 ms bins during the cue epoch. Bar plots show the averages and standard deviations of the mean for the y-axis. The r values represent the Spearman correlations. **A**-**C**, The vertical position of each dot represents the cross-validated accuracy of a classifier that predicts the reward probability condition from the neural activity (see Methods). The dashed line represents chance accuracy. **D**-**F**, The vertical position of the dots represents the ω_p^2 reward probability effect size calculated in 100 ms (**C** and **D**) and 20 ms bins (**E** and **F**) in the cue epoch. **A**, **D** and **G** show neurons from the vermis. **B**, **E** and **H** show neurons from the basal ganglia.



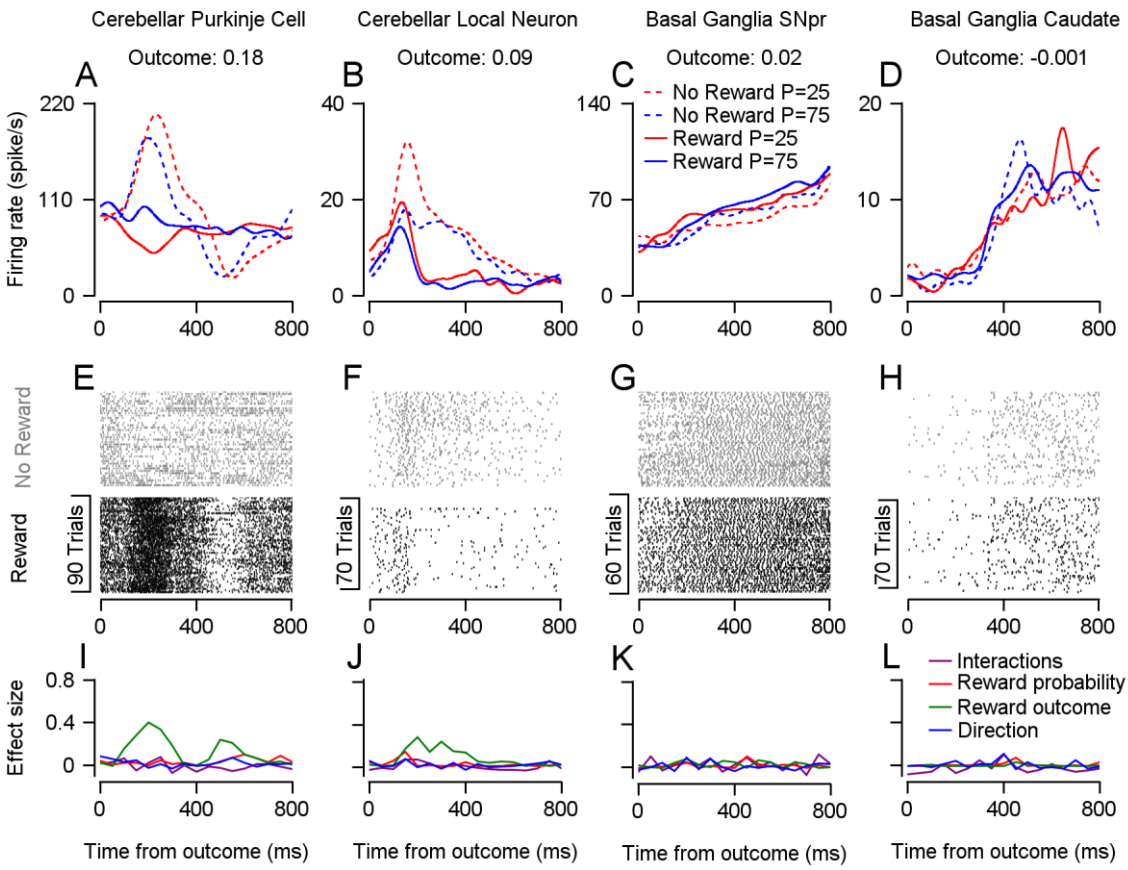
<u>Extended Data Fig. 3:</u> ω_p^2 effect sizes are not determined by firing rate. Each dot represents the firing rate (horizontal) during a specific epoch and effect size (vertical) for a single neuron. **A** and **B** show the reward probability effect size in the cue epoch. **C** and **D** show the direction effect size in the target motion epoch. **E** and **F** show the reward outcome effect size during the outcome epoch. **A**, **C** and **E** show the cerebellum populations and **B**, **D** and **F** the basal ganglia. The r values represent the Spearman correlations.



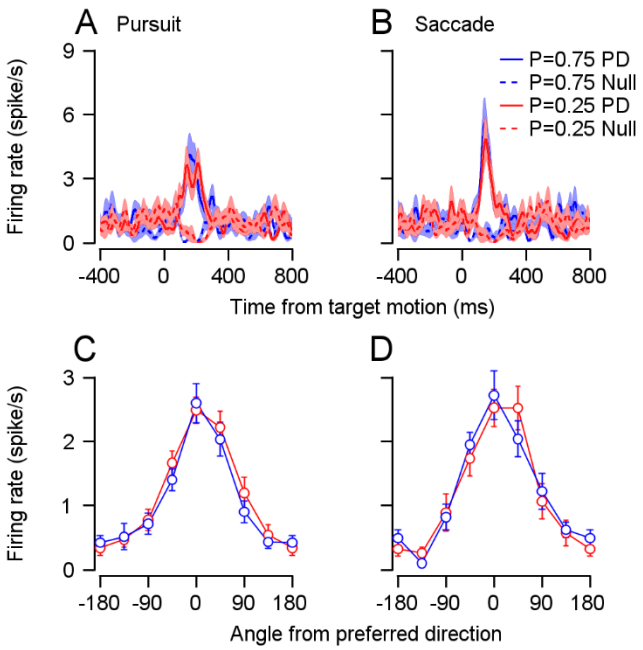
<u>Extended Data Fig. 4:</u> Licking behavior is similar between the reward probability conditions during the cue epoch. A and **B**, Fraction of trials with licks during the cue epoch in each probability condition for monkeys A (**A**) and G (**B**).



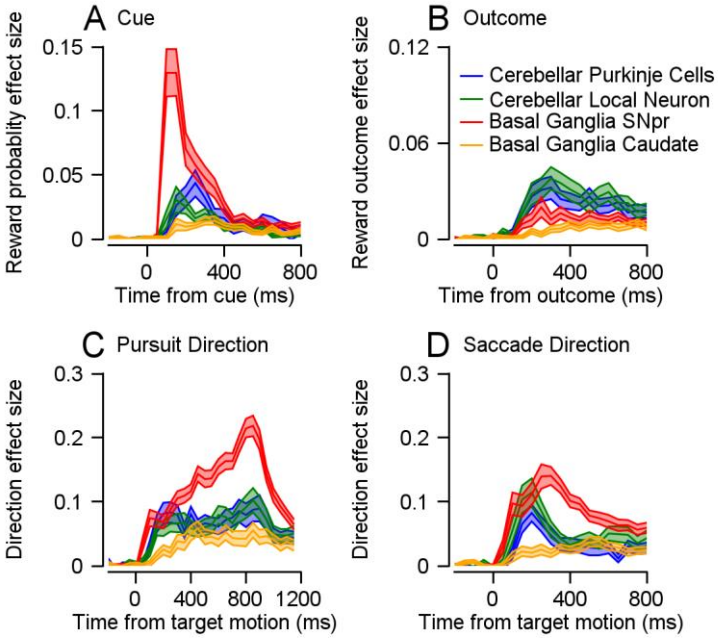
<u>Extended Data Fig. 5:</u> Examples of responses to target motion in the different populations. A-D, PSTHs aligned to the onset of the target motion for the different direction conditions. **A** shows a Purkinje cell on the saccade task, **B** a local cerebellar neuron on the pursuit task, **C** a caudate neuron on the pursuit task, and **D** basal ganglia SNpr neuron on the pursuit task. The entire trial direction effect sizes are shown above the PSTHs. **E-H**, Raster plots for the same neurons. **I-L**, Direction (blue), reward probability (red), and reward probability x direction interaction (purple) effect sizes for the same neurons.



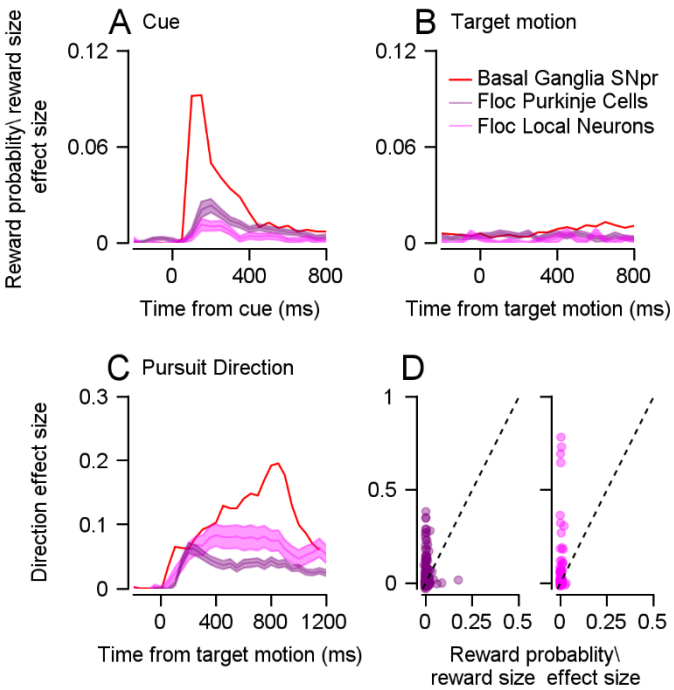
<u>Extended Data Fig. 6:</u> Examples of responses in the outcome epoch in the different populations. A-D, PSTHs aligned to the end of the trial for the different reward probabilities and reward outcome conditions. **A** shows a Purkinje cell, **B** a local cerebellar neuron, **C** a basal ganglia SNpr neuron and **D** a caudate neuron. **E-H**, Raster plots for the same neurons. **I-L**, Direction (blue), reward probability (red), reward outcome (green), and interactions (purple) effect sizes for the same neurons.



<u>Extended Data Fig. 7:</u> The complex spike response to direction is not modulated by the reward probability condition. A and **B**, Average complex spike PSTHs in the preferred direction (PD, solid) and the direction 180° to it (null, dashed) in the P=0.25 (red) and P=0.75 (blue) conditions in the pursuit (**A**) and saccade (**B**) tasks. **C** and **D**, Average complex spike tuning curves aligned to the preferred direction of each neuron in the pursuit (**C**) and saccade (**D**) tasks for the different reward conditions. These analyses were performed on 15 significantly directionally tuned complex spikes on the saccade task, and 17 on the pursuit task. Significant tuning was defined using a rank-sum test on the complex spike rate 100 to 300 ms following the target motion onset. Error bars represent the standard error of the mean over individual neurons.



<u>Extended Data Fig. 8:</u> Effect sizes of neurons with a significant time-varying response. A-D, This figure depicts a replication of the analysis presented in Fig. 2A (**A**), 4C (**B**), 3A (**C**), and 3B (**D**), for the subsets of neurons that had a significant time-varying response in the corresponding epoch (see Methods). Repeating the statistical tests did not alter the conclusions reported in the main text. Test for **A**: Permutation Welch's ANOVA test: $p < 0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}purkinle}$, $p_{SNpr\text{-}local\ cerebellar} < 0.001$. Test for **B:** Permutation Welch's ANOVA test: $p < 0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}cerebellum}$, $p_{caudate\text{-}cerebellum} < 0.001$, $p_{Purkinje\text{-}cerebellar\ local} = 0.33$. Test for **C**: Permutation Welch's ANOVA test: $p < 0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}purkinle} < 0.001$, $p_{SNpr\text{-}local\ cerebellar} = 0.46$. Test for **D**: Permutation Welch's ANOVA test: $p < 0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}purkinle}$, $p_{SNpr\text{-}local\ cerebellar} < 0.001$.



<u>Extended Data Fig. 9:</u> Effect sizes of neurons in the floccular complex of the cerebellum and neighboring areas during pursuit. Effect sizes were calculated for neurons collected from two other monkeys in a previous study[1,2]. The monkeys performed a similar pursuit task in which the cue indicated the future reward size. Pink traces represent Purkinje cells in the floccular complex and purple traces represent local neurons. Red traces are the average effect size of the SNpr shown previously. **A** and **B**, Average reward size or reward probability effect size in the cue (**A**; Permutation Welch's t-test: $p_{SNpr\text{-}purkinje}$, $p_{SNpr\text{-}local\ cerebellar}<0.01$) and target motion (**B**; Permutation Welch's t-test: $p_{SNpr\text{-}purkinje}$, $p_{SNpr\text{-}local\ cerebellar}<0.001$). **C**, Average direction effect size in the target motion epoch (Permutation Welch's t-test: $p_{SNpr\text{-}purkinje}<0.001$, $p_{SNpr\text{-}local\ cerebellar}=0.05$). **D**, Each dot represents a single neuron's reward probability (horizontal) and direction (vertical) effect sizes in the target motion epoch (Bootstrap t-test: $p_{purkinje}$, $p_{local\ cerebellar}<0.01$).