Organization of reward and movement signals in the basal ganglia and cerebellum

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 is thought to be involved in adjusting sensorimotor parameters. Recent findings of reward signals in the cerebellum have challenged this dichotomous view. To compare the basal ganglia and the cerebellum directly, 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 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 findings suggest that reward and movement information is sharpened through the basal ganglia, resulting in a higher signal-to-noise ratio than in the cerebellum.

Introduction

One of the core roles of the brain is to generate movements that increase the chance of gaining rewards. To achieve this, sensory information along with predictions about the reward outcomes of actions is processed into movement. This process is dispersed across multiple brain structures. To understand how reward is used to drive movement, the contribution of each area to the transformation of reward information into a motor command needs to be disentangled. One possible approach 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]. Thus, a system approach that compares signals across populations is better positioned to yield insights into the transformations of signals as they propagate across neural populations.

The basal ganglia and cerebellum are two major subcortical structures that control movement. The basal ganglia are considered to be involved in choices between actions and reward-based learning[4,5]. By contrast, the cerebellum has been hypothesized to be involved in adjusting movement and error-based motor learning[4,6,7]. However, recent findings of reward signals in the cerebellum have challenged this dichotomous division of labor [1,8–12], and imply that the fundamental difference between these structures may not be so straightforward[13,14]. Although reward signals in the basal ganglia and cerebellum exhibit some similarity, they currently cannot be directly compared, since experiments have been conducted on different behaviors, tasks, and species, and in parts of these structures that are involved in the production of different movements.

To overcome the limitations of previous research, we compared signals in the basal ganglia and the cerebellum directly by recording from oculomotor areas in both structures in the same tasks and monkeys. Focusing on oculomotor regions allowed us to build on well-established knowledge on the coding of eye movements[15–17], the connectivity and anatomy of the oculomotor system[18], and the causal relationship between structures and movements[19–23] to study how reward and movement signals are multiplexed in these structures. 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). The activity of the caudate and the SNpr neurons is modulated by both eye movements and reward, thus supporting the claim that the basal ganglia implement reward-based learning[17,24–26]. In the cerebellum, we recorded from the oculomotor vermis which is causally involved in the production of saccade and pursuit eye movements[20,21]. The vermis has also been suggested to participate in cognitive and affective processing[27–29], making it an intriguing region of the cerebellum to study the interactions between reward and motor signals. We designed tasks that manipulated eye movement direction and the probability of receiving a reward, which enabled us to examine reward expectations, motor signals, and how they interact.

We found that reward expectation and eye movement direction signals were the most pronounced in the output of the basal ganglia, intermediate in the vermis, and the smallest in the input of the basal ganglia. This suggests that information is sharpened through the basal ganglia to a greater extent than through the cerebellar cortex. These results extend recent findings of reward coding in the cerebellum and provide the first direct comparison with the basal ganglia.

Results

Eye movement tasks that manipulate reward probability and movement direction

To compare the coding of eye movements and reward directly in the vermis and the basal ganglia, we recorded from both areas in the same monkeys (Fig. S1). 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[30] and the monkey was required to smoothly track the target (Fig. 1B top, Fig. S2). In the saccade task, the target disappeared and reappeared at one of eight 10° eccentric locations. Shortly (221±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, Fig. S2). At the end of a successful trial, the target disappeared, and the monkey received a reward with the probability specified by the target color.

Before recording neural activity, we investigated the relationship between reward probability and eye movements by examining the behavior in additional tasks. We chose reward probabilities of 0.25 and 0.75, since they generated substantial reward expectation differences that were easily noticeable by the monkeys. We confirmed this through a separate task, in which the monkeys were required to track one of two targets based on reward probability associations (see Methods, Fig. 1C). Although there were differences in behavior between conditions in both the saccade and pursuit tasks, these differences were very small (see Methods, Fig. S2), thus allowing us to disambiguate the coding of motor variables from the coding of reward.

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.

<u>Direct comparison of neurons in the basal ganglia and vermis</u>

To examine how reward and movements are represented across structures, we recorded from areas in the basal ganglia and cerebellum involved in the generation of pursuit and saccades[15–17,31–33]. In the basal ganglia, we recorded from the body of the caudate, an input structure, and the SNpr, an output structure. 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[34,35]. Other neurons were classified as putative local neurons that do not project outside the cerebellar cortex.

In each population, we found that some neurons responded to the presentation of the reward probability cue (Fig. 1D-G, Fig. S3). To compare these responses to the cue between populations, we calculated the reward probability effect size. Specifically, we used the partial ω^2 (ω_p^2, 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[36] (Fig. 1H). 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 within conditions, whereas values close to 0 indicate smaller or no differences between conditions. For example, the example SNpr neuron (Fig. 1F) had consistent responses within conditions that differed markedly between reward probability conditions by either pausing or increasing its rate in the low and high probability conditions. In contrast, the caudate, vermis Purkinje cell, and the local vermis local neuron (Fig. 1D, E, and G) exhibited less consistent responses, resulting in lower ω_p^2. These examples demonstrate that the effect size reflects the signal-to-noise ratio in 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 the reward probability was high (for example, Fig. 1E and F) while others had higher rates when the 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 (Fig. S3)[37,38]. Second, the variance explained by additional variables is removed from the noise term. This is critical when comparing populations that multiplex the coding of experimental variables to different extents. 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, in contrast to other effect size measures in which 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 example neuron. All neurons showed an increase in effect size after the appearance of the cue when the reward probability information first became available. However, this increase was substantially larger in the SNpr neuron. This increase corresponded to the time within the trial when the difference between the probability condition in the SNpr neuron was large and the within-condition variability was small. Intuitively, a large effect size is correlated with improved decoding of the experimental condition from the neural activity (Fig. S4A-C).

We also calculated the effect size over the entire trial (see Methods). This yielded a single assessment of the reward probability coding throughout the entire epoch. These whole epoch effect sizes are specified in the caption of Fig. 1D-G. We verified that the effect sizes were robust to different bin sizes (Fig. S4D) and that they were not determined by the firing rates of the neurons (Fig. S5).

<u>The coding of reward probability is more pronounced in the SNpr than in the vermis 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 each population (Fig. 2A). We pooled the 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 the 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, there was a transient increase in the reward probability effect size in all populations. This effect was the most pronounced in the SNpr (Permutation Welch's ANOVA test: $p<0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}purkinje}$, $p_{SNpr\text{-}local}<0.001$). 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. 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 (Fig. S6).

This analysis confirms recent findings reporting reward-related activity in the cerebellum[1,9–11] and the numerous studies indicating reward-related activity in the basal ganglia[38–40]. However, beyond

these studies, these findings demonstrate that the SNpr responses were substantially stronger than the other populations. The results indicate a strong sharpening of reward signals within the basal ganglia, from the caudate to the SNpr. The finding that the caudate exhibited a smaller average effect size than both vermis populations is surprising given the hypothesis that the basal ganglia are the primary site where reward signals coincide with the coding of movements[4].

The coding of eye movement direction is more pronounced in the SNpr than in the vermis and the caudate

To study how eye movement signals distribute across subcortical networks we next compared the coding of eye movement directions between populations on the pursuit and saccade tasks (Fig. S7). Some neurons in our sample had complex temporal dynamics[41] (for example, Fig. S7C). The ω_p^2 effect size analysis allowed us to compare the coding of directions between tasks and populations without averaging in time to calculate the tuning curve or having to define the preferred direction of neurons[42]. We therefore calculated the direction effect size of each population and task. All the populations exhibited an increase in direction effect sizes as the target began to move, on both the pursuit and the saccade tasks (Fig. 3A and B). This increase was the most pronounced in the SNpr (Pursuit – Permutation Welch's ANOVA test: $p<0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}Purkinje}$, $p_{SNpr\text{-}local}<0.001$; Saccades - Permutation Welch's ANOVA test: $p<0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}Purkinje}<0.001$, $p_{SNpr\text{-}local}=0.15$). On the saccade task, the difference between the SNpr and local vermis 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, since the vermis is thought to be functionally closer to behavior[20,23].

Common control for saccades and pursuit

A substantial fraction of neurons were recorded on both the pursuit and saccades tasks (see Methods). We used ω_p^2 to revisit the question of common control for saccade and pursuit[43] in the vermis[20] and the SNpr[33] and examined this question for the first time in the caudate. We found significant correlations between the direction effect sizes on the pursuit and saccade tasks in all four populations (Fig. 3D and E; Spearman correlation: $r_{SNpr}=0.5$, $r_{caudate}=0.5$, $r_{Purkinje}=0.66$, $r_{local}=0.65$, $p<0.001$ 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 claims of common control for saccades and pursuit in the subcortical networks.

Coding of reward probability during movement

Our analysis method allowed us to quantify the extent to which reward was coded during movement. One of the main benefits of ω_p^2 is that it takes the variance in the response of neurons that is related to additional experimental variables into account. The noise term in ω_p^2 is defined such that variance that can be explained by experimental variables is not treated as noise (Fig. 1H). This allows us to compare the coding of reward in populations that encode the direction of eye movements to varying extents. When the targets began to move, the reward probability effect size was small in all populations (Fig. 4A). SNpr neurons showed a relatively late increase in the coding of reward probability which was not observed in any of the other populations. 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. 4B, most

dots fall above the identity line, Bootstrap t-test: $p<0.001$ for all populations, see examples in Fig. S7I-L). These results are consistent with previous studies on the basal ganglia[44], cerebellum[11], and frontal eye field[45], 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.

While the coding of the reward condition itself was slight, reward expectation may modulate the coding of direction, having an indirect effect on the neural activity. Neurons in which the reward probability condition substantially modulates the coding of direction should display high effect size values for reward probability and direction interaction (reward probability x direction). These modulations can arise from sharpening or gain modulations of the tuning curve[46]. We did not find such interactions in any of the populations (Figure 4C), suggesting that at least on these tasks, reward and direction coding have a minimal influence on one another (Bootstrap t-test: $p_{SNpr}=0.18$, $p_{caudate}=0.41$, $p_{Purkinje}=0.35$, $p_{local}=0.32$). Thus, the separation in the timing of the coding of reward and movement and the lack of interaction during movement indicate that these areas multiplex reward and movement mostly across task epochs rather than simultaneously. The coding of reward during movement is thought to reflect reward bias in behavior[25]. However, our tasks were designed to minimize behavioral differences between reward probability conditions (Fig. S2) so that only a small portion of the activity in the dataset would be associated with the reward modulation of behavior.

<u>The coding of reward outcome is comparable in the vermis and the basal ganglia</u>

Reward outcome and prediction error signals were found in the basal ganglia[5] and recently in the cerebellum[1,9,47,48]. Here, we used the trials in which the reward was delivered or omitted to compare the reward outcome signals in the basal ganglia and the cerebellum directly. 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. We pooled neurons across the pursuit and saccade tasks since in the outcome epoch, the tasks are similar.

We calculated the ω_p^2 to partition the responses into reward outcome, reward probability, and direction during the outcome epoch (examples in Fig. S8). 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. 5A; Bootstrap t-test: $p<0.01$, in all populations). The average direction effect sizes of the population preserved the same order as when aligned to target motion (Fig. 5B; Permutation Welch's ANOVA test: $p<0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}Purkinje}$, $p_{SNpr\text{-}local}<0.01$). Coding of direction was expected since monkeys typically saccade back from the different eccentric positions at the end of a trial (saccade latency from target disappearance: 316±191 ms for Monkey A, 407±107 ms for Monkey G). Thus, finding the largest coding of movement during the outcome epoch in the SNpr supports our conclusion that the coding of eye movement direction is the most pronounced in the output of the basal ganglia. We found that the average reward outcome effect size was slightly higher in the vermis than in the caudate and the SNpr (Fig. 5C; Permutation Welch's ANOVA test: $p<0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}vermis}$, $p_{caudate\text{-}vermis}<0.001$, $p_{Purkinje\text{-}local}=0.57$). This result differs considerably from the coding of reward probability during the cue (Fig. 3A) in which the SNpr exhibited the largest response.

We cannot completely rule out the possibility that some of the activity in this epoch was related to food consumption or other movements after reward delivery and the end of the trial. We verified that our results could not be explained by blinks and saccades that occur when the target disappeared by removing trials with blinks and equating the number of saccades across the reward

223 outcome conditions (see Methods). We found that in these conditions the effect sizes remained
224 significantly positive for all populations (Bootstrap t-test: $p<0.001$ for all populations).

225 According to the reward prediction error hypothesis, reward predictability should modulate
226 the response to the reward outcome[49–52]. In terms of effect size, prediction error or any outcome
227 modulations based on reward predictability (such as signed or unsigned prediction errors) should be
228 expressed in a positive reward probability x reward outcome interaction effect size. However, we
229 found only a slight interaction in all populations (Fig. 5D; Bootstrap t-test: p_{SNpr}, $p_{local}<0.01$,
230 $p_{Purkinje}=0.06$,$p_{caudate}=0.14$). The coding of reward outcome was stronger than the coding of their
231 interaction (Fig. 5E and F; Bootstrap t-test: $p<0.01$ in all populations) indicating that reward prediction
232 error coding and any other interaction between probability and outcome were at best only weakly
233 coded in the recorded populations. This suggests a different functional role for the recorded
234 populations as compared to dopaminergic neurons[2] or climbing fibers[8,10] that code the reward
235 prediction error.

236 In our tasks, the monkeys were not explicitly signaled when the reward was omitted, which
237 may have caused the reward prediction signals to be spread out in time and could have impeded our
238 ability to detect them. However, the end of each trial was clearly marked by the disappearance of the
239 colored target which enabled the monkeys to infer whether they would receive a reward shortly after
240 the trial ended. We did not find differences in neural activity between rewarded trials with different
241 reward probabilities (P=0.25 reward vs. P=0.75 reward), as expressed in the effect sizes for the
242 comparison between these two conditions (not shown). In these trials, the sound of the reward pump
243 and the availability of the reward constituted explicit indications for timing. Thus, the lack of reward
244 prediction signals at the outcome epoch was unlikely to be the outcome as a lack of an explicit timing
245 signal.

246 <u>The coding of reward and movement in vermis complex spikes is small compared to simple spikes</u>

247 Purkinje cells have the unique property of displaying two distinct types of spikes: Simple and
248 complex spikes. Simple spikes are similar to spikes observed in other neurons and were the focus of
249 our analysis thus far. Complex spikes occur at a much lower frequency (~1Hz, see diamonds in Fig. 1F)
250 and are hypothesized to serve as teaching signals that affect plasticity[6,7]. Recently, complex spikes
251 have been shown to code for reward signals[8–11]. We therefore extended our analysis to complex spikes
252 recorded in the Purkinje cell population.

253 The complex spike reward probability effect sizes during the cue epoch (Fig. 6A; Bootstrap t-
254 test: $p<0.05$) and reward outcome during the outcome epoch (Fig. 6B; Bootstrap t-test: $p<0.05$) were
255 significant but small. During the target motion epoch, the effect size of the reward was not significantly
256 different from 0, consistent with previous results[11] (not shown; Bootstrap t-test: $p=0.55$). The coding
257 of direction during the motion epoch on the saccade (Fig.6C; Bootstrap t-test: $p<0.05$) and pursuit (Fig.
258 6D; Bootstrap t-test: $p<0.01$) tasks was also small relative to all other populations. A paired
259 comparison of the effect sizes of the complex and simple spikes from the same neuron indicated that
260 for all task variables, the effect size was substantially larger for the simple spikes (scatter plots in Fig.
261 5A-D, Bootstrap t-test: $p<0.05$ in all comparisons).

262 Note that effect size characterizes the signal-to-noise ratio of single neurons. Thus, a large
263 unaccounted variance relative to the effect will result in a small effect size. We confirmed that the
264 population response, when averaged over trials and neurons, was modulated by the direction of

movement (Fig. S9). 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 constrains their ability to transform information at the single neuron level. While complex spikes may serve as a teaching signal that promotes plasticity[6,7] in the cerebellum, our analysis indicates that such a teaching signal is very noisy at the level of single neurons.

The coding of eye movement direction in the local vermis population is more pronounced in neurons with short waveforms and low firing rates.

The cerebellar cortex contains multiple cell types that do not project outside the cerebellar cortex but are involved in computation. To investigate whether different cell types could represent eye movement directions to different extents, we plotted the direction effect size of local vermis neurons along with the width of their waveforms (see Methods) and their average firing rates. We found that local vermis neurons with narrow waveforms and low firing rates tended to have higher direction effect sizes (Fig. 6E, Permutation Welch's ANOVA test: $p<0.01$, for the four groups separated by the median values of the firing rate and waveform width), indicating that specific subtypes of local neurons may code for direction to different extents. We repeated the same analysis for Purkinje cells but did not find this trend (Fig. 6F, Permutation Welch's ANOVA test: $p=0.43$, for the four groups separated by the median values of the firing rate and waveform width). Previous research has suggested that the lower firing rate neurons with narrow waveforms are mossy fibers,[53] implying that inputs to the vermis are more strongly tuned than interneurons or vermis output. However, since we did not find distinct clusters in our data, we refrain from making strong claims about specific populations.

Controlling for sampling bias

One of the challenges when comparing coding across populations or structures is that the populations themselves may not be homogenous. Specifically, some sub-populations may be over or under-sampled. 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 oculomotor areas within the basal ganglia and cerebellum. Thus, we compared subsets of neurons that were 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 with neurons that showed a time-varying response to the task (see Methods, Fig. S10); namely, neurons whose firing rate was significantly different across the different time bins in an epoch.

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 (Fig. S11). In a previous report, we examined the coding of reward size and movement in the floccular complex and its neighboring areas[11] in a task similar to the pursuit task in which the color of the cue indicated the future reward size instead of its probability (see Methods). We reanalyzed the floccular complex data and found that the reward effect sizes in the floccular complex were comparable to the vermis reward probability effect sizes. 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

We recorded from the basal ganglia and vermis of the same monkeys as they performed saccade and pursuit eye movement tasks in which we manipulated reward expectation. Recording from the same monkeys enabled us to control for differences between individuals and draw direct comparisons across structures. We found similar results on the saccade and pursuit tasks, which suggests that our findings are not limited to the specific parameters of our tasks, but may reflect the general properties of the eye movement system.

Our results extend recent findings on reward signals in the cerebellum[1,8–11] and relate them directly to signals found in the basal ganglia[5,24,39]. In comparison to the basal ganglia output structure, the reward expectancy (Fig. 2) and sensorimotor (Fig. 3) signals in the vermis were small. Reward outcome signals in the vermis were comparable to the basal ganglia (Fig. 5), suggesting that the reward consequences of behavior are processed in both areas similarly. The findings also showed that reward outcome signals were not modulated by reward predictability (Fig. 5D), 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[54–58]. 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 that eye movement direction should be more strongly represented in downstream areas. However, 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[18,20,23] 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 for the basal ganglia and the vermis (Fig. 5).

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[59]. Thus, in the basal ganglia, noisy representations across many neurons are transformed into a smaller population 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[59,60]. 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[61].

In the cerebellum, we only found a small difference in the extent to which local vermis neurons and Purkinje cells coded for reward and movement information, despite the vast convergence onto Purkinje neurons. One possibility is that the convergence of local neurons onto Purkinje neurons may not lead to an increase in the signal-to-noise ratio. This could be due to the averaging out of opposing signals or high noise correlations limiting the cancellation of correlated noise[62–64]. While diverse inputs onto Purkinje cells may cancel each other out, they may be advantageous during the learning process by providing multiple connections that facilitate the rapid acquisition of new associations or the adjustment of movements[65]. Alternatively, the signal-to-noise ratio of local neurons and Purkinje cells might reflect recurrent connectivity within the cerebellar cortex rather than feedforward processing. For example, Purkinje cell collaterals projecting back to the cerebellar cortex[65–69] along with the

353 projections from local neurons to the Purkinje cells may cause local neurons to code for similar
354 information as Purkinje cells, highlighting the importance of Purkinje cell collaterals back to the
355 cerebellar cortex[66–70]. Further research is needed to disentangle these two possibilities.

356 Our study focused on the oculomotor regions of the basal ganglia and cerebellum, which
357 enabled a direct comparison of reward and movement signals in areas that contribute to the same
358 behavior. We presented data from the input or local populations (caudate and local vermis neurons)
359 as well as the output populations (SNpr and Purkinje cells) in both the basal ganglia and cerebellar
360 cortex. We observed greater sharpening of the signals in the basal ganglia than in the cerebellum.
361 However, due to their specific circuitries, defining homologies between these structures is challenging.
362 A full comparison between the regions needs to include the entirety of the basal ganglia nuclei, the
363 cerebellar cortex, and the deep cerebellar nuclei. In particular, the globus pallidus externus nucleus in
364 the basal ganglia, an additional input to the SNpr, presents an interesting target for investigation, given
365 the observed low signal-to-noise ratio in the caudate relative to the SNpr. The small effect sizes in the
366 Purkinje cells underscore the importance of investigating the deep cerebellar nuclei. This type of
367 exploration would help indicate whether a low signal-to-noise ratio is characteristic of the entire
368 cerebellar network or whether there are functional distinctions between the cerebellar cortex and
369 nuclei.

370 Using effect size measures, we suggested a framework to compare populations that is agnostic
371 to the direction of the response, takes trial-by-trial variability into account, allows for comparisons of
372 specific variables while controlling for others, and can be used to examine the interactions between
373 variables. This framework can serve to investigate additional regions in both structures where eye
374 movement signals have been observed[71,72] as well as in regions involved in the production of other
375 movements. For instance, in the cerebellum, reward signals have been observed in areas associated
376 with arm movements[73]. Similarly, in the basal ganglia, reward signals have been reported in various
377 contexts and regions[74]. Thus, our study can form the basis for revising our understanding of the roles
378 played by the basal ganglia and the cerebellum in the integration of reward and motor signals.

379 **Methods**

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

389 Visual stimuli were displayed on a monitor (55 and 63 cm from the eyes of monkeys A and G).
390 The stimuli appeared on dark background in a dimly lit room. A computer performed all real-time
391 operations and controlled the sequences of target motions. The position of the eye was measured
392 with a high temporal resolution camera (1 kHz, Eyelink 1000 plus, SR research) and collected for
393 further analysis. We performed two subsequent surgeries to place a round 19 mm diameter recording
394 cylinder over the basal ganglia and the vermis. For each structure, up to 5 quartz-insulated tungsten
395 electrodes (impedance of 1-2 Mohm) were lowered through each cylinder using the Mini-Matrix
396 System (Thomas Recording GmbH). When lowering the electrodes, we searched for neurons that

397 responded during pursuit or saccade eye movements, but sometimes collected data from neurons
398 that did not respond to eye movements.

399 We began by recording from the caudate and then moved to the SNpr, while continuously
400 recording from the vermis. To record from the caudate we approximated its depth relative to the
401 recording chamber based on an MRI scan (Fig. S1). We then lowered electrodes towards the target.
402 We verified that we had reached caudate by identifying typical tonically active neurons with broad
403 waveforms and a low and regular firing rate (~5 Hz)[75]. Based on the extensive mapping of the caudate
404 and the MRI, we estimated the location of the SNpr relative to the recording chamber. We lowered
405 the electrodes to this location. At the targeted site we identified neurons with a high baseline firing
406 rate and the typical extracellular shape of SNpr neurons[76]. We confirmed that some of these neurons
407 exhibited a clear pause during saccades in certain directions. On some recording days, we also
408 identified neurons with pronounced eye position sensitivity and a regular firing rate, as expected from
409 the neighboring third oculomotor nerve. To locate the oculomotor vermis, we lowered electrodes
410 through the tentorium towards the vermis and identified areas with characteristic background
411 responses to saccades[15]. We classified neurons as Purkinje cells by the presence of complex spikes[77,78].

412 Table 1 shows the number of neurons recorded from each population. We only included
413 neurons that were recorded for at least 55 successful trials in either the saccade or the pursuit task,
414 with a median of 141 trials. In the analyses of the cue and outcome epochs, we pooled neurons across
415 tasks. If a neuron was recorded in more than one task, we included the task in which the neuron was
416 recorded for more trials.

	Pursuit task	Saccade task	At least one task
Caudate	86/75	62/107	109/123
SNpr	62/104	38/112	66/165
Local cerebellar neurons	61/42	37/50	72/62
Purkinje cell simple spikes	39/32	29/36	47/43
Purkinje cell complex spikes	56/29	35/38	64/46
Simple and complex spike pairs	36/23	27/26	38/29

417 **Table 1**: Number of neurons recorded from each monkey on each task. The left and right values show
418 the number of neurons for Monkeys A and G.

419 Signals were digitized at a sampling rate of 40 kHz (OmniPlex, Plexon). For the detailed data
420 analysis, we sorted spikes offline (Plexon). For sorting, we used principal component analysis and
421 corrected manually for errors. In some of the Purkinje cells, the complex spikes had distinct low-
422 frequency components. In these cells, we used low-frequency features to identify and sort the
423 complex spikes[77,78]. We paid special attention to the isolation of spikes from single neurons. We
424 visually inspected the waveforms in the principal component space and only included neurons for
425 further analysis when they formed distinct clusters. The sorted spikes were converted into time-
426 stamps with a time resolution of 1 ms and were inspected again visually to check for instability and
427 obvious sorting errors.

428 We also recorded licking behavior to control for behavioral differences between reward

429 conditions that might confound the results. Licks were recorded using an infra-red beam. Monkey A
430 tended not to extend its tongue, therefore we recorded its lip movements.

431 Experimental design

432 Before beginning the neural recordings, we examined the behavior of the monkeys under
433 different reward probability conditions (Fig. S2). Our selection of probabilities of 0.25 and 0.75 for
434 recordings has several advantages. First, the considerable difference between probabilities prompted
435 a noticeable difference in reward expectation before delivery. On an additional task in which the
436 monkeys were presented with two targets moving in orthogonal directions and were rewarded based
437 on the target they tracked, the monkeys consistently chose to track the 0.75 probability target (Fig.
438 1C, see the Choice task section below for more details) indicating that they associated the color of the
439 target with the probability of reward. Second, a probabilistic reward made it possible to test for
440 differences in neural activity between reward delivery and omission (reward outcome) and
441 manipulate reward expectation at delivery to test whether the neural activity was consistent with the
442 reward prediction error hypothesis[49]. Third, the effect of the reward probability on pursuit and
443 saccades was small in comparison to a condition in which the monkeys never received a reward (P=0,
444 Fig. S2), and in comparison, to the variability in the behavior within conditions. This similarity in
445 behavior served to examine the responses to reward information separately from movement
446 preparation and execution.

447 **Pursuit task:**

448 Each trial started with a bright white target that appeared in the center of the screen (Fig. 1A).
449 After 500 ms of presentation, in which the monkey was required to acquire fixation, a colored target
450 replaced the fixation target. The color of the target signaled the probability of receiving a reward upon
451 successful tracking of the target. For monkey A we used yellow to indicate a 0.75 probability for reward
452 and green to indicate 0.25. For monkey G we reversed the associations. After a variable delay of 800-
453 1200 ms, the targets stepped in one of eight directions (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°) and
454 then continuously moved in the opposite direction (step-ramp)[30]. For both monkeys, we used a target
455 motion of 20°/s and a step to a position 4° from the center of the screen. The target moved for 750
456 ms for monkey A and 650 ms for monkey G and then stopped and stayed still for an additional 500-
457 700 ms. If the monkey's gaze was within a 3-5°x3-5° window around the target, the monkey received
458 a reward according to the probability specified by the color.

459 **Saccade task:**

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

465 **Choice task:**

466 We used this task to verify that the monkeys associated the color of the target with the reward
467 probability. The structure of the task was similar to previous reports[79]. The monkeys were required to
468 choose one of two targets with different reward probability values presented on the screen. Their
469 choice determined the probability of receiving a reward. Each trial began with a 500 ms fixation

period, similar to the tasks described previously. Then two additional colored spots appeared at a location eccentric to the fixation target. One of the colored targets appeared 4° below or above the fixation target (vertical axis) and the other appeared 4° to the right or left of the fixation target (horizontal axis). We use the same colors as we used in the saccade and pursuit tasks. The monkeys were required to continue fixating on the fixation target in the middle of the screen. After a variable delay of 800-1200 ms, the white target disappeared, and the colored targets started to move towards the center of the screen (vertically or horizontally) at a constant velocity of 20°/s. After a variable delay, the monkeys typically made saccades toward one of the targets. We defined these saccades as an eye velocity that exceeded 80°/s. The target that was closer to the endpoint of the saccade remained in motion for up to 750 ms and the more distant target disappeared. The monkeys were required to track the target until the end of the trial. After the end of the motion, the remaining target stayed still for an additional 500-700 ms, and then the monkeys received a reward according to the probability determined by the color of the chosen target. To allow the monkeys to choose the pursuit target freely, we suspended fixation requirements from the time of disappearance of the fixation target to the end of the first saccade. After the saccade, the monkeys' gaze had to stay within a 4-5°x4-5° window around the target. Upon successful performance, the monkeys received a reward with the probability defined by the color of the target it tracked.

Reward size task:

We used previously analyzed data recorded from the floccular complex and adjacent areas, while the monkeys performed a smooth pursuit task in which we manipulated reward size. The full details can be found in previous reports[11,80]. Briefly, the temporal and target motion properties of the task were the same as the reward probability pursuit task described above. However, in this task, the color of the target signaled the size of the reward the monkeys would receive if they tracked the target. One color was associated with a large reward (~0.2ml) and the other with a small reward (~0.05 ml).

Data analysis

All analyses were performed using MATLAB (MathWorks).

Eye movement analysis:

To calculate the average of the smooth pursuit initiation, we first removed the saccades and treated them as missing data. 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. We then averaged the traces in the target movement direction. Finally, we smoothed the traces using a Gaussian filter with a standard deviation of 20 ms.

Peristimulus time histogram:

To examine 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 20 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 (Fig. S9), 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.

Waveform width:

The waveform width was calculated by first aligning the waveforms to their minimal points, and then calculating the time difference between the peak and the trough.

Effect size in time:

To quantify a neuron's coding of a variable over 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[36,81,82]. 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 epoch 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 the Type II sum of squares to account for unequal group sizes (for example, a smaller number of trials in the reward-delivered as compared to the reward-omitted within the P=0.25 condition)[83]. 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 calculated 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 (Fig. S4D-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:

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

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 (Fig. S5). Although Purkinje cells had the highest firing rates, their effect sizes during the cue and target motion epochs were relatively small. Additionally, in some cases, the rate and effect sizes were correlated, but this correlation was small or absent in others.

Time-varying response:

We repeated our analysis on subsets of neurons within each population that changed their average firing rates significantly throughout the epoch (Fig. S10). To find these neurons we calculated the significance of the time variable, using the same ANOVA design described above in the Effect Size subsection over the entire epoch.

	Cue	motion: pursuit	motion: saccade	Outcome
Purkinje cell simple spikes	48 (0.53)	49 (0.69)	45 (0.69)	86 (0.95)
Local cerebellar neurons	78 (0.58)	70 (0.67)	66 (0.75)	118 (0.88)
SNpr	162 (0.7)	146 (0.87)	123 (0.82)	213 (0.92)
Caudate	92 (0.39)	71 (0.44)	104 (0.61)	171 (0.73)

Table 2: Number and fraction (in parenthesis) of neurons that were included in the analysis presented in Fig. S10.

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 (Fig. S4). 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:

The distributions of effect sizes tended to be non-normal and the variances were different between 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, since 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 (two-tailed test).

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 (two-tailed test).

Blink and saccade control during the outcome epoch:

Monkeys tended to blink at the end of trials, immediately after the target disappeared and the reward had either been delivered or omitted. The blink frequency was slightly higher at the end of trials when the reward was omitted than when it was delivered (a difference of approximately 3 Hz). The monkeys also tended to saccade back towards the center of the screen at the end of trials, with a slightly higher frequency of early saccades when a reward was delivered than when it was omitted (a difference of approximately 2 Hz). To confirm that the reward outcome effect size did not result from differences in the pattern of saccades or blinks across task conditions, we repeated the analysis after removing trials with a blink in the first 500 ms, and equated the number of saccades between reward conditions by randomly discarding trials. Even after this manipulation, the effect sizes were significantly positive in all populations. This suggests that the reward outcome effect persisted regardless of the occurrence of blinks and early saccades.

Additional Information

Data availability: The data that support the findings of this study are available from the authors upon reasonable request (mati.joshua@mail.huji.ac.il).

Code availability: All code used in this study is available on GitHub (https://github.com/noga-larry).

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 Israel Science Foundation (380/17).

Author Contributions

The experiment was designed by all three authors. The data were collected by NL and GZ, who recorded from the vermis and basal ganglia, respectively. The data were analyzed by NL and GZ. The manuscript was written by NL and MJ, and all authors were involved in the interpretation of the results.

Competing Interests

The authors declare no competing interests.

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. Kameda, M., Niikawa, K., Uematsu, A. & Tanaka, M. Sensory and motor representations of internalized rhythms in the cerebellum and basal ganglia. *Proceedings of the National Academy of Sciences* **120**, (2023).

15. 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).

16. 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).

17. 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).

18. 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).

19. Keller, E. L., Slakey, D. P. & Crandall, W. F. Microstimulation of the primate cerebellar vermis during saccadic eye movements. *Brain Res* **288**, 131–143 (1983).

20. 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).

21. Noda, H. & Fujikado, T. Topography of the oculomotor area of the cerebellar vermis in macaques as determined by microstimulation. *J Neurophysiol* **58**, 359–378 (1987).

22. Nakamura, K. & Hikosaka, O. Facilitation of Saccadic Eye Movements by Postsaccadic Electrical Stimulation in the Primate Caudate. *Journal of Neuroscience* **26**, 12885–12895 (2006).

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

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

25. Kawagoe, R., Takikawa, Y. & Hikosaka, O. Reward-Predicting Activity of Dopamine and Caudate Neurons - A Possible Mechanism of Motivational Control of Saccadic Eye Movement. *J Neurophysiol* **91**, 1013–1024 (2004).

26. Hikosaka, O., Sakamoto, M. & Usui, S. Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *https://doi.org/10.1152/jn.1989.61.4.814* **61**, 814–832 (1989).

27. Schmahmann, J. D. An Emerging Concept: The Cerebellar Contribution to Higher Function. *Arch Neurol* **48**, 1178–1187 (1991).

28. Schmahmann, J. D. The role of the cerebellum in affect and psychosis. *J Neurolinguistics* **13**, 189–214 (2000).

29. Guell, X., Schmahmann, J. D., Gabrieli, J. D. E. & Ghosh, S. S. Functional gradients of the cerebellum. *Elife* **7**, (2018).

30. Rashbass, C. The relationship between saccadic and smooth tracking eye movements. *J Physiol* **159**, 326–338 (1961).

31. 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).

32. 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).

33. 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).

34. 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).

35. 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).

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

37. 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).

38. 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).

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

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

41. Zur, G., Larry, N. & Joshua, M. High-Dimensional Encoding of Movement by Single Neurons in Basal Ganglia Output. *bioRxiv* 2023.05.17.541090 (2023) doi:10.1101/2023.05.17.541090.

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

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

44. 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.

45. 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).

46. 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).

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

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

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

50. 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).

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

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

53. Prsa, M., Dash, S., Catz, N., Dicke, P. W. & Thier, P. Characteristics of responses of Golgi cells and mossy fibers to eye saccades and saccadic adaptation recorded from the posterior vermis of the cerebellum. *Journal of Neuroscience* **29**, 250–262 (2009).

54. 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).

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

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

57. 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).

58. 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).

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

60. 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).

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

62. Zohary, E., Shadlen, M. N. & Newsome, W. T. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**, 140–143 (1994).

63. Moreno-Bote, R. *et al.* Information-limiting correlations. *Nat Neurosci* **17**, 1410–1417 (2014).

64. Larry, N. & Joshua, M. Neural correlations are attenuated in the transformation from the frontal cortex to movement. *bioRxiv* 2021.04.29.441941 (2021).

65. 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).

66. Crook, J. D., Hendrickson, A., Erickson, A., Possin, D. & Robinson, F. R. Purkinje Cell Axon Collaterals Terminate on Cat-301+ Neurons in Macaca Monkey Cerebellum. *Neuroscience* **149**, 834 (2007).

67. Hirono, M. *et al.* Cerebellar globular cells receive monoaminergic excitation and monosynaptic inhibition from Purkinje cells. *PLoS One* **7**, (2012).

68. Guo, C., Rudolph, S., Neuwirth, M. E. & Regehr, W. G. Purkinje cell outputs selectively inhibit a subset of unipolar brush cells in the input layer of the cerebellar cortex. *Elife* **10**, (2021).

69. Witter, L., Rudolph, S., Pressler, R. T., Lahlaf, S. I. & Regehr, W. G. Purkinje Cell Collaterals Enable Output Signals from the Cerebellar Cortex to Feed Back to Purkinje Cells and Interneurons. *Neuron* **91**, 312–319 (2016).

70. Guo, C., Rudolph, S., Neuwirth, M. E. & Regehr, W. G. Purkinje cell outputs selectively inhibit a subset of unipolar brush cells in the input layer of the cerebellar cortex. *Elife* **10**, (2021).

71. Ashmore, R. C. & Sommer, M. A. Delay activity of saccade-related neurons in the caudal dentate nucleus of the macaque cerebellum. *J Neurophysiol* **109**, 2129–2144 (2013).

72. 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).

73. Sendhilnathan, N., Goldberg, M. E. & Ipata, A. E. Mixed Selectivity in the Cerebellar Purkinje-Cell Response during Visuomotor Association Learning. *Journal of Neuroscience* **42**, 3847–3855 (2022).

74. Schultz, W. Reward functions of the basal ganglia. *Journal of Neural Transmission* vol. 123 679–693 Preprint at https://doi.org/10.1007/s00702-016-1510-0 (2016).

75. Joshua, M., Adler, A., Mitelman, R., Vaadia, E. & Bergman, H. Midbrain Dopaminergic Neurons and Striatal Cholinergic Interneurons Encode the Difference between Reward and Aversive Events at Different Epochs of Probabilistic Classical Conditioning Trials. *Journal of Neuroscience* **28**, 11673–11684 (2008).

76. Schultz, W. Activity of pars reticulata neurons of monkey substantia nigra in relation to motor, sensory, and complex events. *J Neurophysiol* **55**, 660–677 (1986).

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

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

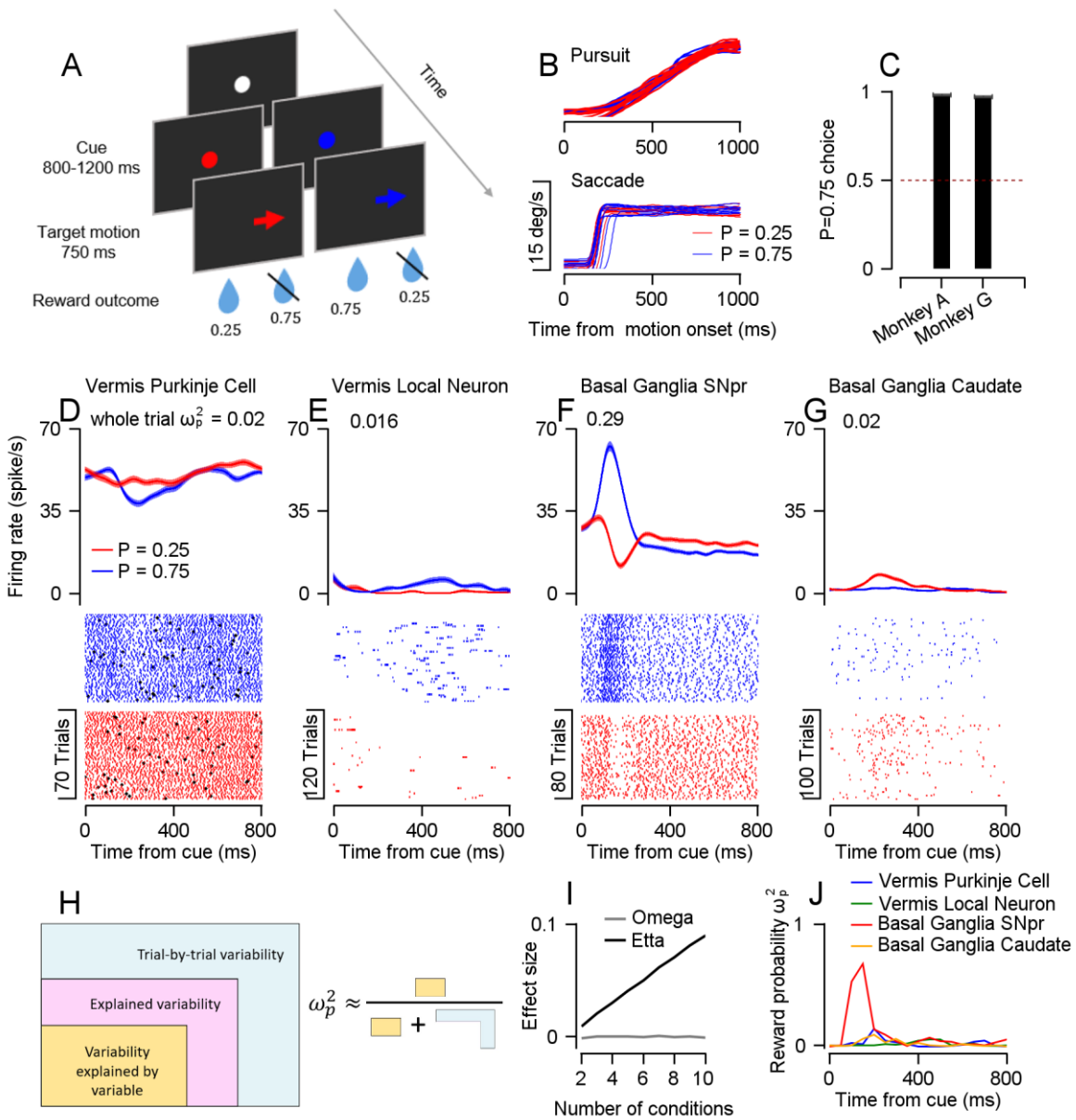
79. Joshua, M. & Lisberger, S. G. Reward Action in the Initiation of Smooth Pursuit Eye Movements. *Journal of Neuroscience* **32**, 2856–2867 (2012).

80. 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).

81. 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).

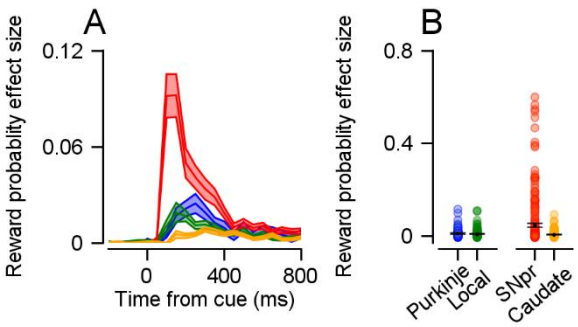
82. Okada, K. Is Omega Squared Less Biased? a Comparison of Three Major Effect Size Indices in One-Way Anova. *Behaviormetrika* **40**, 129–147 (2013).

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

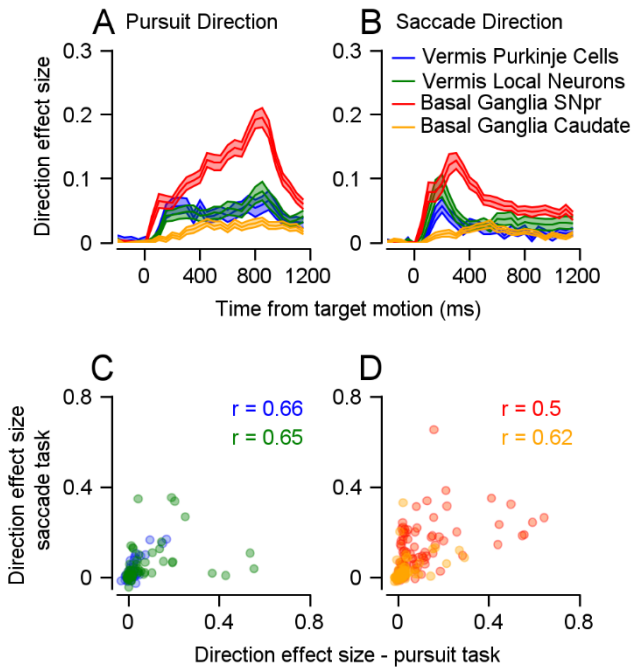


<u>Fig. 1:</u> Effect size analysis to quantify the coding of reward probability on a trial-by-trial basis. A, Schematics of the pursuit and saccade eye movement tasks. The spots represent the target and the arrows show the direction of continuous motion (pursuit) or instantaneous jump of the target (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 vermis Purkinje cell (**D**), vermis local neuron (**E**), SNpr neuron (**F**), and caudate neuron (**G**). The black diamonds in **D** represent complex spikes recorded in the same Purkinje cell. Error bars on PSTHs represent SEMs over trials. **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

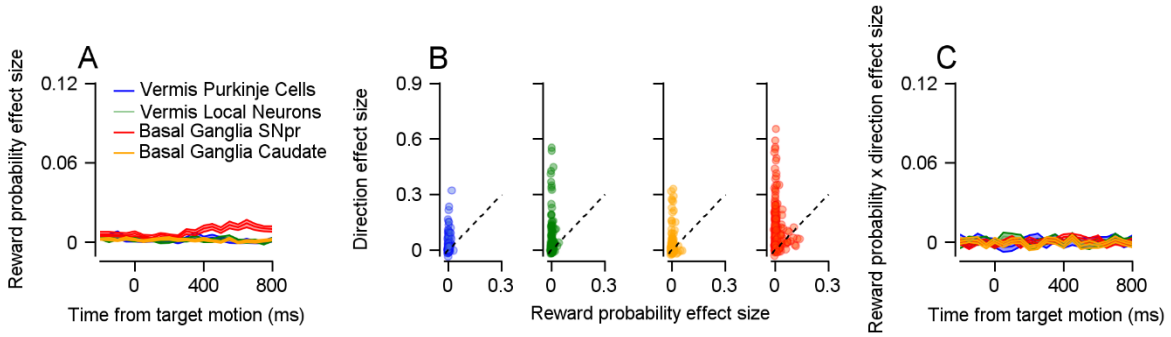
(black) effect sizes as a function of the number of conditions in a simulation in which the neural responses were independent of the condition. **J**, Reward probability ω_p^2 effect sizes of neurons in **D**-**G** calculated in bins of 50 ms. Entire trial effect sizes that represent the total coding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



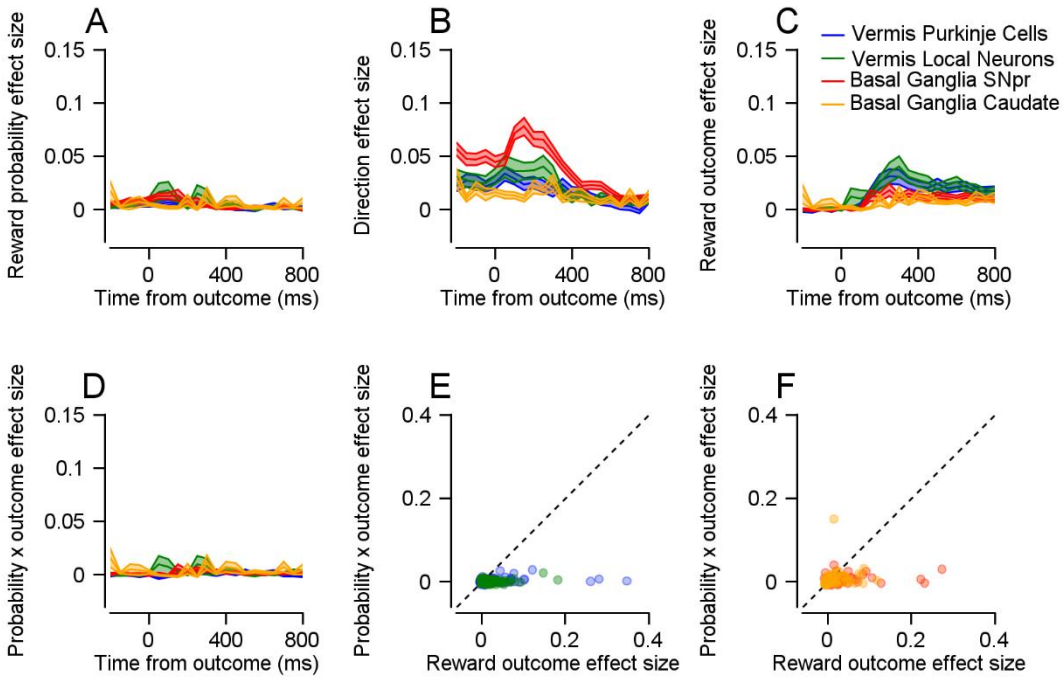
<u>Fig. 2:</u> Coding of reward probability during the cue epoch. A, Average reward probability effect size in the cue epoch as a function of time. **B**, Each dot represents a single neuron's reward probability effect size in the entire cue epoch. Bars show the average and standard error of the mean.



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



<u>Fig. 4:</u> Coding of reward probability during the motion epoch. A, Average reward probability effect size in the motion epoch. **B**, Each dot represents a single neuron's reward probability (horizontal) and direction (vertical) effect sizes in the target motion epoch. Dashed lines correspond to the identity line. **C,** Average reward probability x direction interaction effect size in the motion epoch.



<u>Fig. 5:</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.

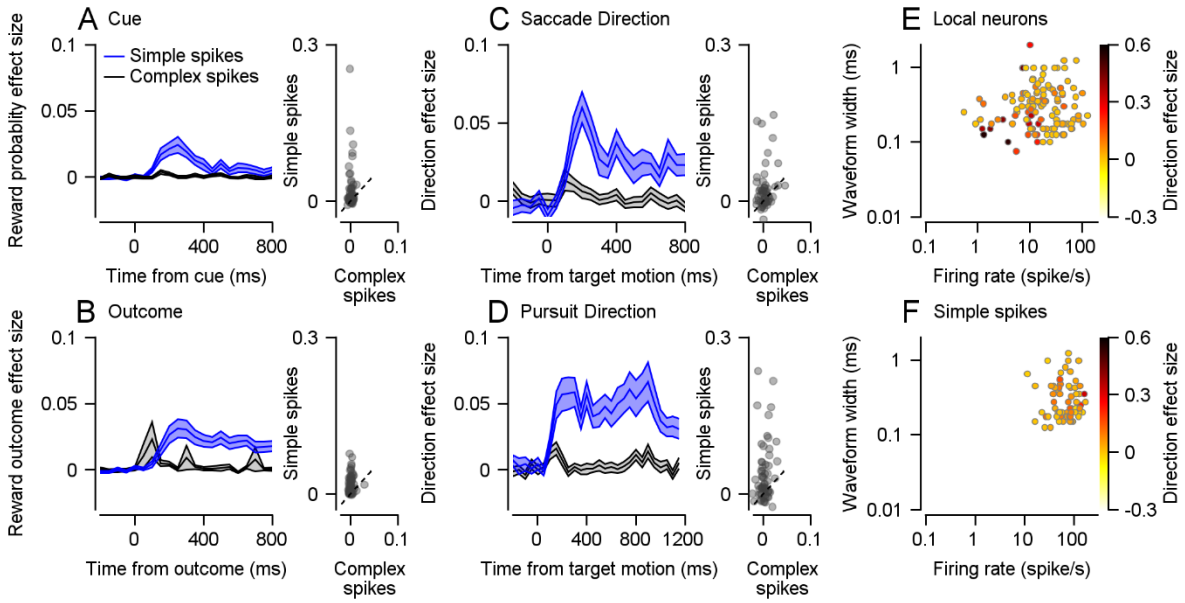
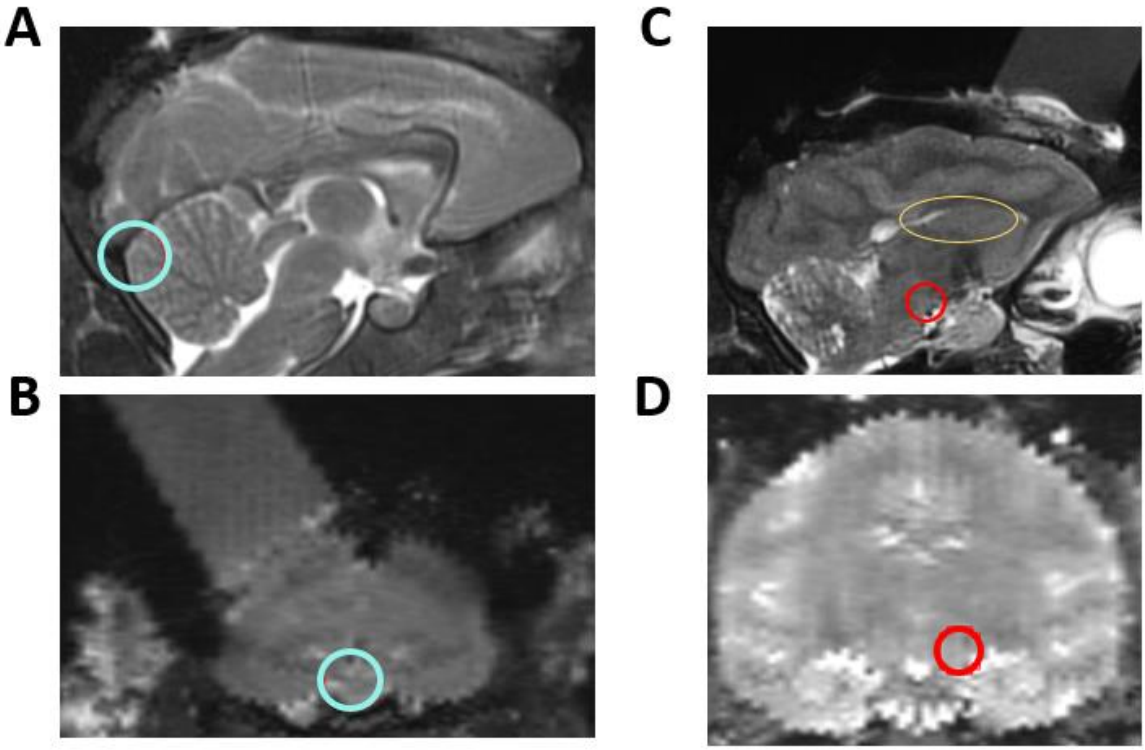
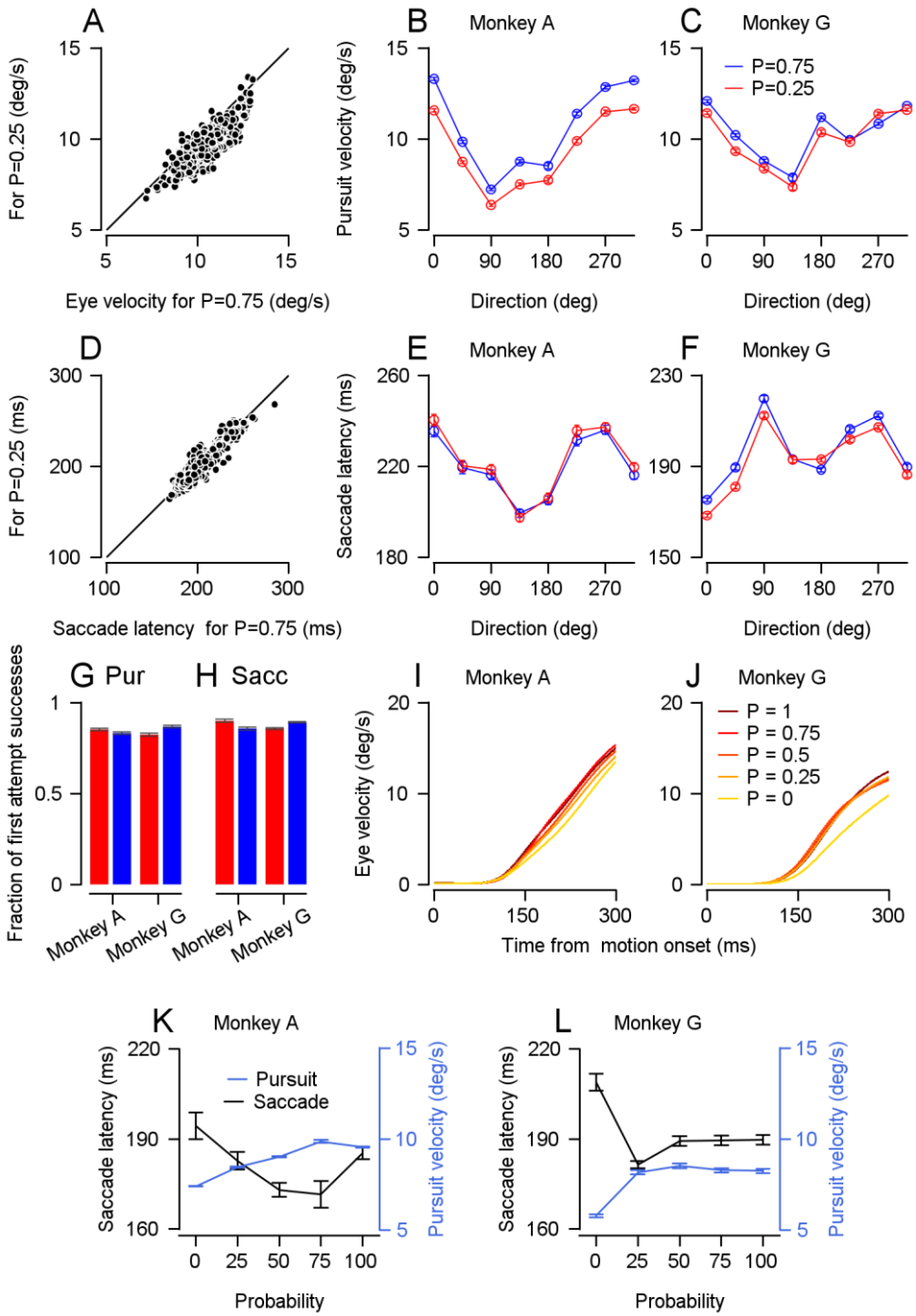


Fig. 6: Effect sizes of the vermis populations. A-D, Black traces represent complex spikes and blue traces show the simple spike effect sizes displayed previously. In the scatter plots, each dot represents a single Purkinje cell's complex (horizontal) and simple (vertical) spike effect sizes. The dashed lines show the identity line. **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. **E** and **F**, each dot represents the firing rate (horizontal) and the waveform width (vertical) of a single vermis local neuron (**E**) or a Purkinje cell's simple spikes (**F**) on a log-log scale. The color represents that neuron's direction effect size in the motion epoch of either the saccade or pursuit task.

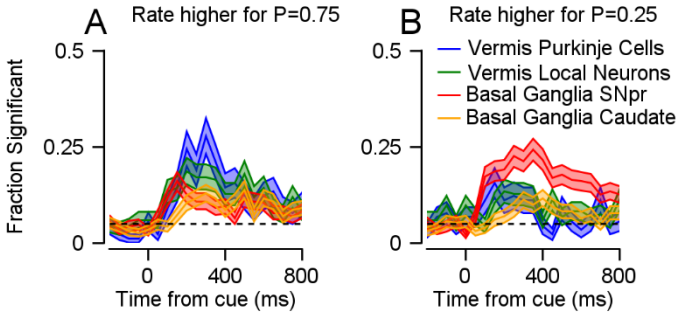


<u>Fig. S1:</u> MRI sections of the recorded structures. A, Mid-sagittal section showing the oculomotor vermis, marked by the cyan circle. **B**, Coronal section, 32 mm posterior to the anterior commissure, showing the recording chamber above the same area as in **A**. **C**, Sagittal section, 5 mm left of the midline, showing recording chamber above the body of the caudate, marked by the yellow ellipse, and the SNpr, marked by the red circle. **D**, Coronal section 7 mm posterior to the anterior commissure, showing the SNpr, marked by the red circle.

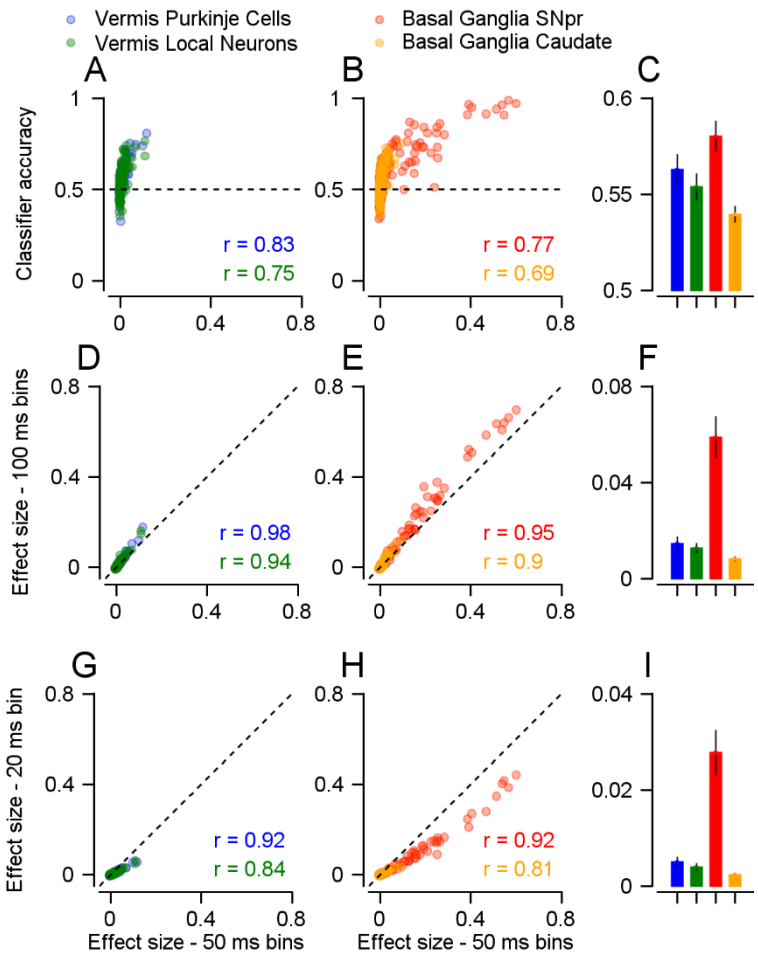


<u>Fig. S2:</u> Pursuit and saccade behavior in the P=0.25 and P=0.75 reward probability conditions, and sessions with five reward probability conditions. A, Each dot represents the average speed for an individual session 200 to 250 ms after target motion onset for the P=0.75 (horizontal) and P=0.25 (vertical) reward probability conditions (Signed-rank: p<0.001, n=501). The line shows the identity line. **B** and **C**, Pursuit velocity 200 to 250 ms after target motion onset as a function of trial direction for monkeys A (**B**) and G (**C**). **D**, Each dot represents the average saccade latency for an individual session (Signed-rank: p<0.001, n=471). **E** and **F**, Saccade latency as a function of trial direction for monkeys A (**E**) and G (**F**). **G** and **H**, Fraction of the trial that the monkey completed on the first attempt in the pursuit task (**G**) and the saccade task (**H**). **I** and **J**, 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 (**I**) and G (**J**). **K** and **L**, Black traces correspond to the vertical axis on the left and show the average saccade

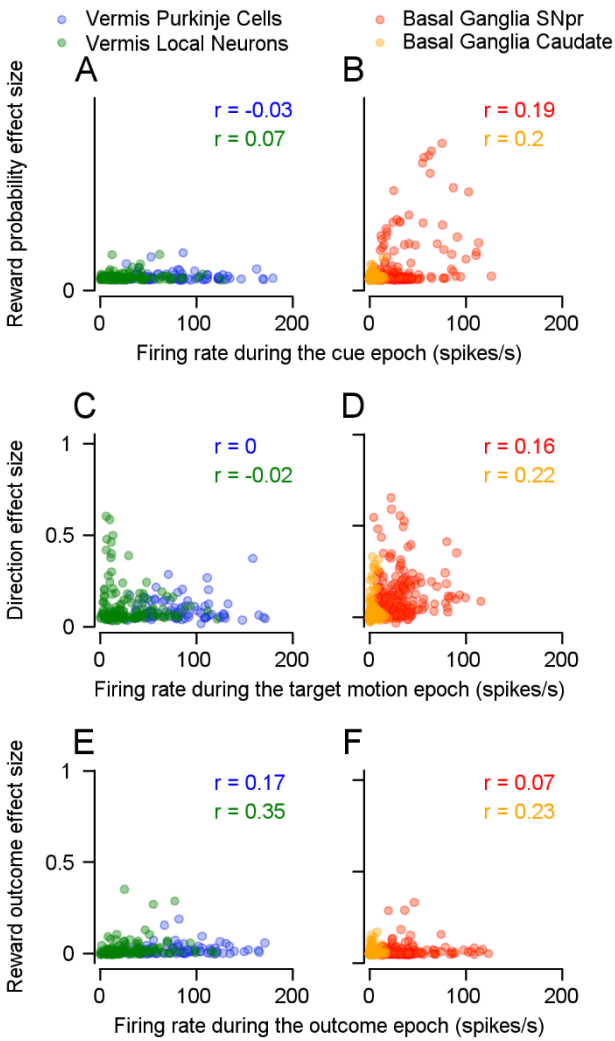
latencies in sessions with multiple probabilities on the saccade task. Blue traces correspond to the vertical axis on the right and show the average pursuit velocity 200 to 250 ms after the target motion in sessions with multiple probabilities in the pursuit task. **K** and **L** show data from monkeys A and G.



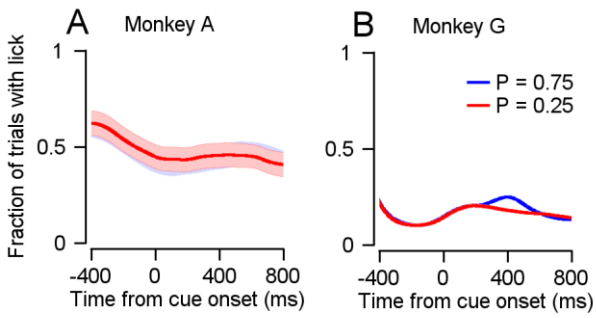
<u>Fig. S3:</u> Fraction of responsive neurons in the cue epoch. A and **B**, The fraction of neurons that had a significantly higher firing rate in the P=0.75 condition (**A**) or the P=0.25 condition (**B**) in 50 ms bins (one-tailed rank-sum test). Dashed lines show the chance level.



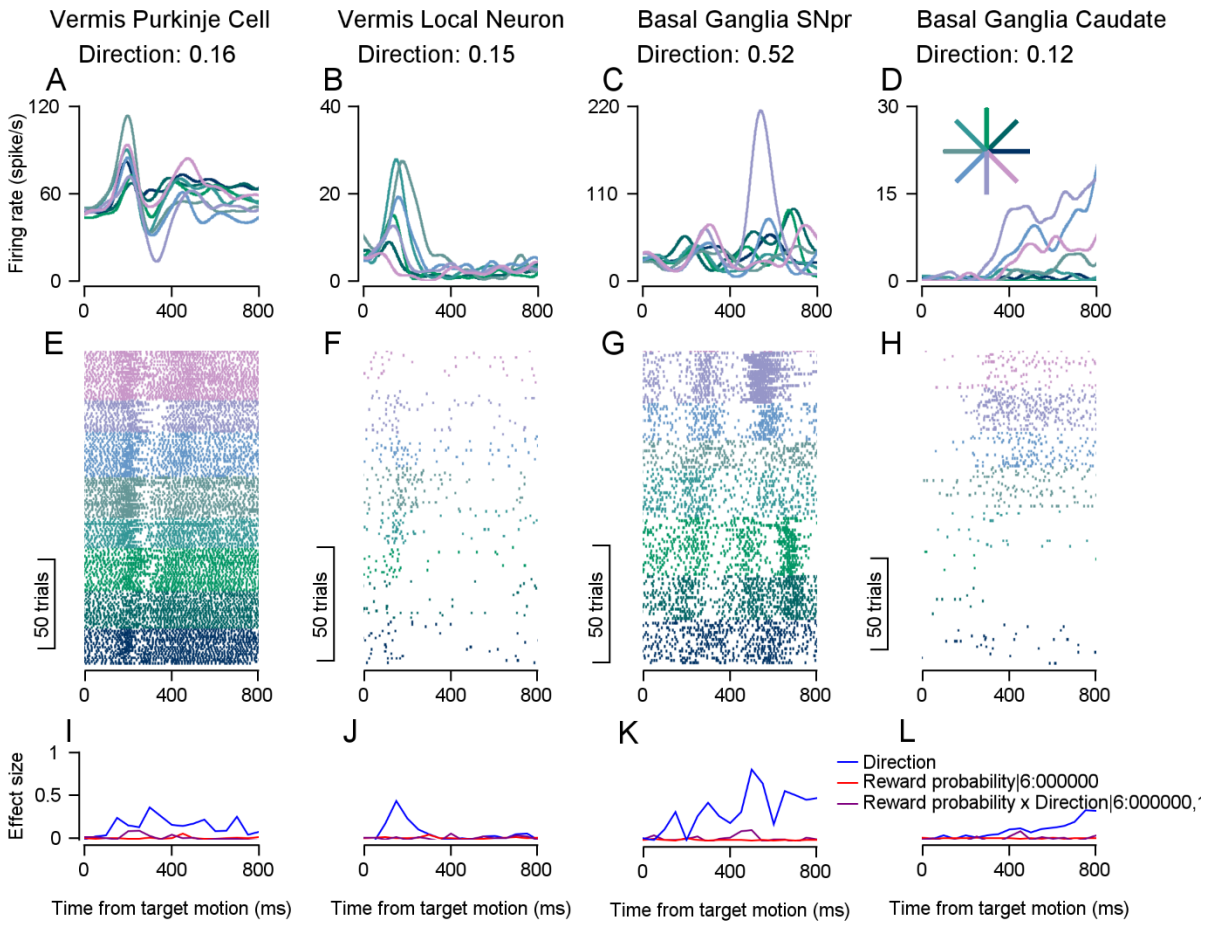
<u>Fig. S4:</u> Correlations of ω_p^2 effect sizes with the accuracy of a classifier and across different bin sizes. Each dot represents a single neuron. In all plots, the horizontal position of the dots represents the ω_p^2 reward probability effect size calculated in 50 ms bins during the cue epoch. Bar plots show the averages and standard deviations of the mean for the y-axis. The r values represent the Spearman correlations (p<0.001 for all 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. The dashed line shows the identity line. **A**, **D**, and **G** show neurons from the vermis. **B**, **E**, and **H** show neurons from the basal ganglia.



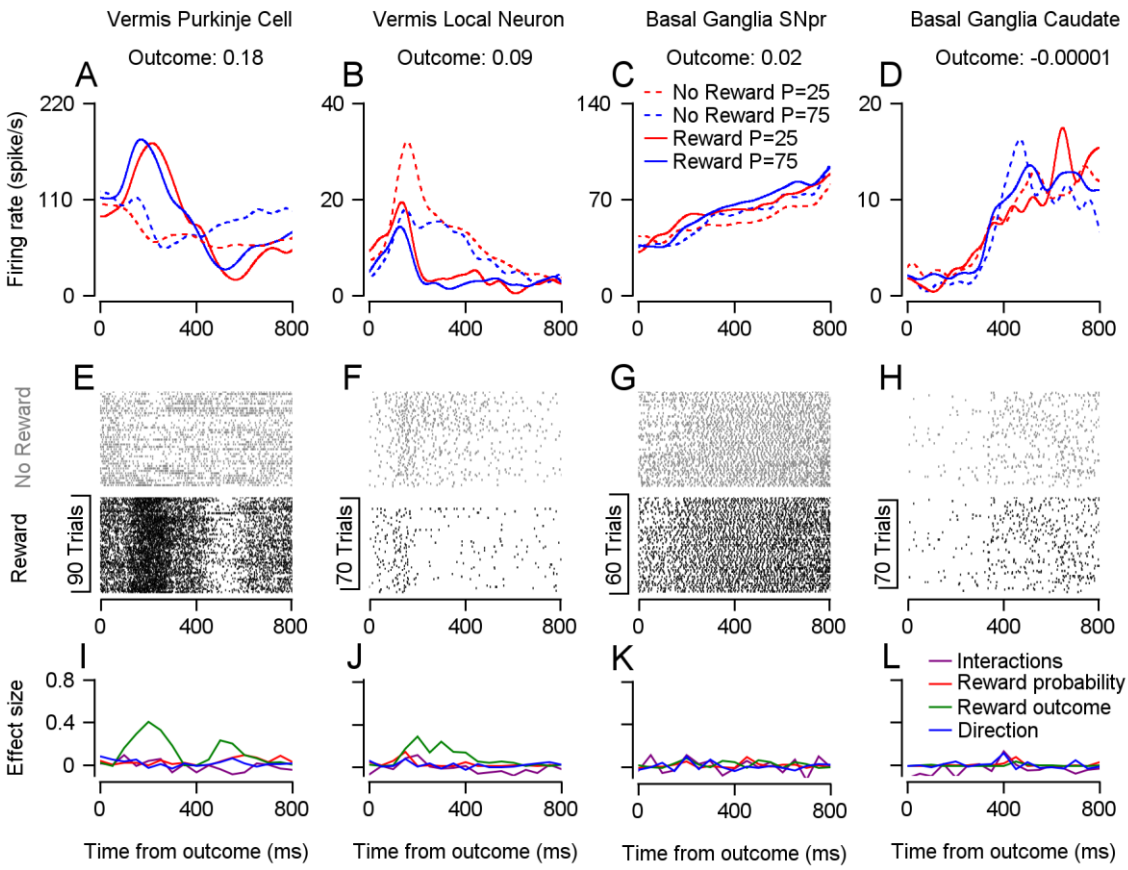
<u>Fig. S5:</u> Correlations between ω_p^2 and firing rate. Each dot represents the firing rate (horizontal) during a specific epoch and the effect size (vertical) for a single neuron. **A** and **B** show the reward probability effect size in the cue epoch. **C** and **D** show the direction effect 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 vermis populations, and **B**, **D,** and **F** the basal ganglia. The r values represent the Spearman correlations.



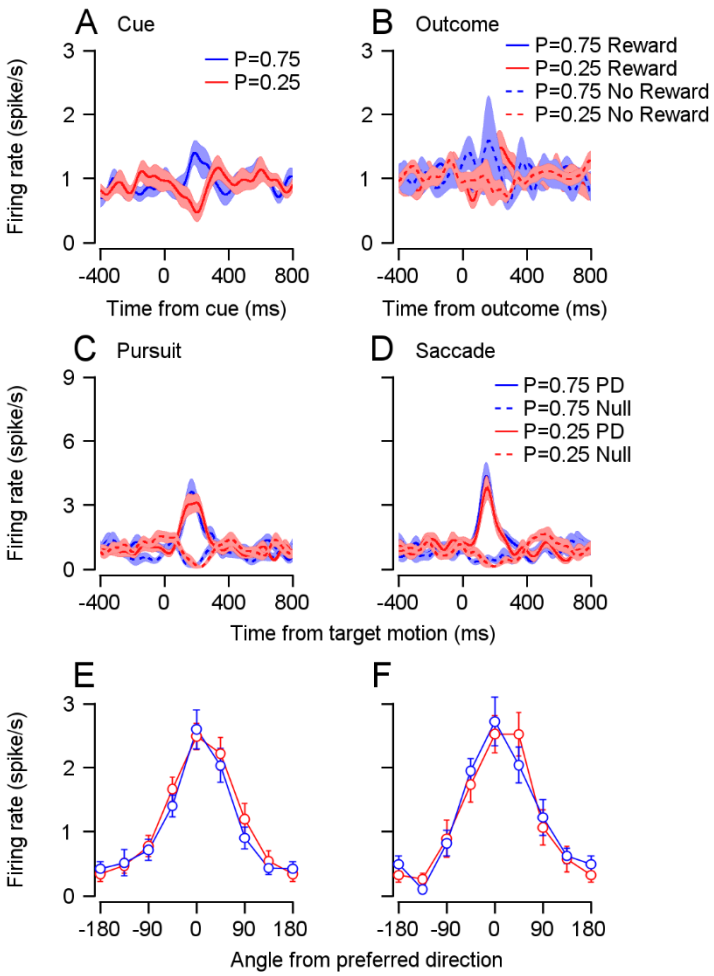
<u>Fig. S6:</u> Licking behavior in the cue epoch. A and **B**, Fraction of trials with licks during the cue epoch in each probability condition for monkeys A (**A**) and G (**B**).



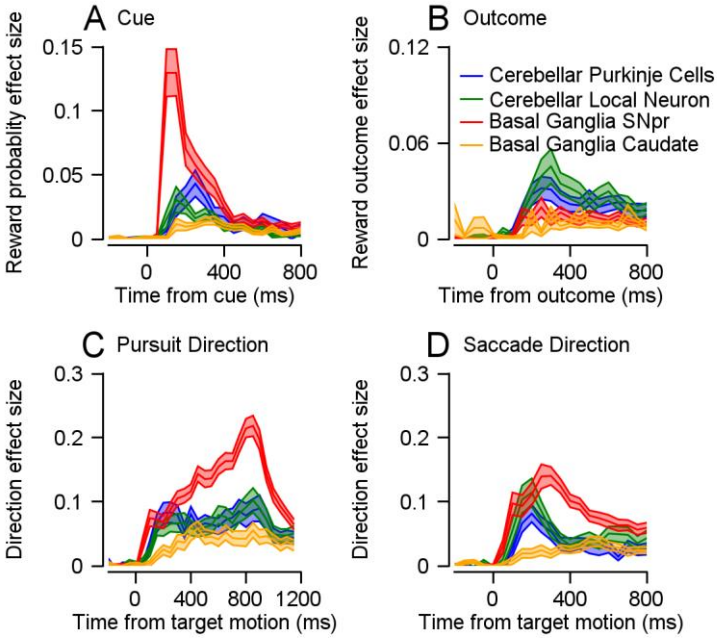
<u>Fig. S7:</u> Examples of responses to target motion in the different populations. A-D, PSTHs aligned to the onset of the target motion for the different direction conditions. **A** shows a Purkinje cell on the saccade task, **B** a vermis local neuron on the pursuit task, **C** a caudate neuron on 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.



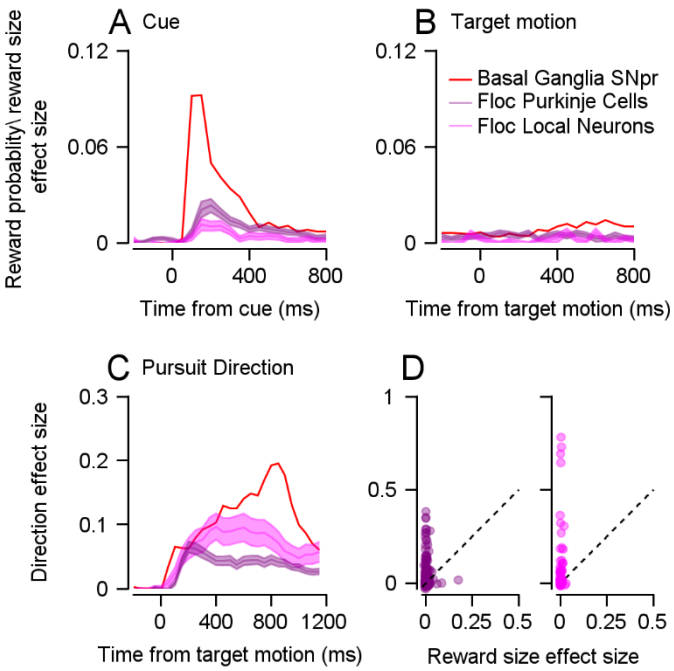
<u>Fig. S8:</u> Examples of responses in the outcome epoch in the different populations. A-D, PSTHs aligned to the end of the trial for the different reward probabilities and reward outcome conditions. **A** shows a Purkinje cell, **B** a vermis local neuron, **C** a basal ganglia SNpr neuron and **D** 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>Fig. S9:</u> Average complex spike responses. A-D. Population complex spike PSTHs during the different task epochs. **A.** Cue epoch, traces show the two reward probability conditions for complex spikes that significantly differentiated between reward probability conditions (ranked sum test on the 100 to 300 ms after the cue onset, 11/110 neurons). **B,** Outcome epoch, traces show different reward probability and reward outcome conditions for complex spikes that significantly differentiated between reward outcome conditions (ranked sum test on the 100 to 300 ms after the reward outcome, 16/110 neurons). **C** and **D**, Motion epoch, traces show 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 (**C**) and saccade (**D**) tasks. **E** and **F**, Average complex spike tuning curves aligned to the preferred direction of each neuron in the pursuit (**E**) and saccade (**F**) tasks for the different reward conditions. **C-F** shows complex spikes that were significantly modulated by direction (Kruskal-Wallis on the 100 to 300 ms after the target motion onset, 15/73 neurons in the saccade task, 17/85 neurons in the pursuit task).



<u>Fig. S10:</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: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}Purkinje}$, $p_{SNpr\text{-}local} < 0.001$. Test for **B:** Permutation Welch's ANOVA test: $p < 0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}vermis}$, $p_{caudate\text{-}vermis} < 0.001$, $p_{Purkinje\text{-}local} = 0.33$. Test for **C**: Permutation Welch's ANOVA test: $p < 0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}Purkinje} < 0.001$. Test for **D**: Permutation Welch's ANOVA test: $p < 0.001$, Permutation Welch's t-test: $p_{SNpr\text{-}caudate}$, $p_{SNpr\text{-}Purkinje} < 0.001$, $p_{SNpr\text{-}local} = 0.41$).



<u>Figure S11:</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: $p_{SNpr\text{-}Purkinje}$, $p_{SNpr\text{-}local}$<0.001) and target motion (**B**; Permutation Welch's t-test: $p_{SNpr\text{-}Purkinje}$, $p_{SNpr\text{-}local}$ <0.001). **C**, Average direction effect size in the target motion epoch (Permutation Welch's t-test: $p_{SNpr\text{-}Purkinje}$<0.001, $p_{SNpr\text{-}local}$=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}<0.01).