

Acquisition of Neural Learning in Cerebellum and Cerebral Cortex for Smooth Pursuit Eye Movements

Jennifer X. Li,^{2,3} Javier F. Medina,⁴ Loren M. Frank,^{2,3} and Stephen G. Lisberger^{1–3}

¹Howard Hughes Medical Institute, ²W. M. Keck Foundation Center for Integrative Neuroscience, and ³Department of Physiology, University of California, San Francisco, San Francisco, California 94143, and ⁴Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

We evaluated the emergence of neural learning in the frontal eye fields (FEF_{SEM}) and the floccular complex of the cerebellum while monkeys learned a precisely timed change in the direction of pursuit eye movement. For each neuron, we measured the time course of changes in neural response across a learning session that comprised at least 100 repetitions of an instructive change in target direction. In both areas, the average population learning curves tracked the behavioral changes with high fidelity, consistent with possible roles in driving learning. However, the learning curves of individual neurons sometimes bore little relation to the smooth, monotonic progression of behavioral learning. In the FEF_{SEM}, neural learning was episodic. For individual neurons, learning appeared at different times during the learning session and sometimes disappeared by the end of the session. Different FEF_{SEM} neurons expressed maximal learning at different times relative to the acquisition of behavioral learning. In the floccular complex, many Purkinje cells acquired learned simple-spike responses according to the same time course as behavioral learning and retained their learned responses throughout the learning session. A minority of Purkinje cells acquired learned responses late in the learning session, after behavioral learning had reached an asymptote. We conclude that learning in single neurons can follow a very different time course from behavioral learning. Both the FEF_{SEM} and the floccular complex contain representations of multiple temporal components of learning, with different neurons contributing to learning at different times during the acquisition of a learned movement.

Introduction

The nervous system can use sensory reports of the errors that occur during one movement to improve the accuracy of subsequent movements, a process commonly referred to as motor learning. Motor learning is inherently dynamic; behavioral improvements occur in increments that become smaller with additional practice. Behavioral learning proceeds continuously and smoothly toward a final asymptote and comprises several smoothly changing components with different rates of learning (Smith et al., 2006; Lee and Schweighofer, 2009; Yang and Lisberger, 2010). Most studies of the neural basis for motor learning have recognized that learning is a circuit-wide process that engages numerous neural components across multiple brain areas (Lisberger, 1994; Boyden et al., 2004; Doyon and Benali, 2005). However, prior work mostly has compared neural responses before versus after learning and has not evaluated the course of acquisition of neural learning. In this study, we recorded the time course of neural learning in multiple brain areas and tested the interesting

possibility that neural learning is episodic in individual neurons and is acquired dynamically across a population of neurons in each area.

Smooth pursuit eye movements offer a number of advantages for analysis of the neural basis of motor learning. If a visual target repeatedly undergoes the same change in direction at a predictable time, then pursuit undergoes reliable learning (Medina et al., 2005). The learned changes in the behavior accrue rapidly and plateau within a few dozen movements. Behavioral experiments have shown that pursuit learning comprises multiple components that emerge over different time courses (Yang and Lisberger, 2010). Pursuit, like most movements, requires contributions from several motor areas, including both the cortex (Keating, 1991; MacAvoy et al., 1991) and the cerebellum (Westheimer and Blair, 1973; Rambold et al., 2002). The portions of the motor cortex and the cerebellum dedicated to pursuit include, respectively, the smooth pursuit region of the frontal eye fields (FEF_{SEM}) and the floccular complex. The FEF_{SEM} is involved in modulating the strength, or gain, of visual–motor transmission (Tanaka and Lisberger, 2001, 2002), and the floccular complex, situated closer to the motor end of the pursuit circuit, explicitly encodes (and determines) the velocity and acceleration of the eye movement. Both of these areas express learned changes in neural responses in association with motor learning in pursuit (Kahlon and Lisberger, 2000; Medina and Lisberger, 2008; Li and Lisberger, 2011).

In this study, we compared the emergence of learned neural activity in the FEF_{SEM} and the floccular complex during the acquisition of the learned behavior. Our findings reveal a mismatch

Received May 19, 2011; revised July 8, 2011; accepted July 13, 2011.

Author contributions: J.X.L., J.F.M., and S.G.L. designed research; J.X.L. and J.F.M. performed research; J.X.L., L.M.F., and S.G.L. analyzed data; J.X.L., J.F.M., L.M.F., and S.G.L. wrote the paper.

This work was supported by the Howard Hughes Medical Institute, National Institute of Mental Health Grant MH077970, and predoctoral fellowships from the National Science Foundation and the National Institutes of Health (NS655982). We thank Drs. Jennifer Raymond, Michael Brainard, and Allison Doupe for helpful comments on a version of the paper. We are grateful to Stefanie Tokiyama, Elizabeth Montgomery, Karen MacLeod, Dirk Kleinhesselink, Scott Ruffner, David Wolfgang-Kimball, Darrell Floyd, and Ken McGary for technical assistance.

Correspondence should be addressed to Dr. Stephen G. Lisberger, Department of Physiology, 513 Parnassus Avenue, Room HSE-802A, San Francisco, CA 94143-0444. E-mail: SGL@phy.ucsf.edu.

DOI:10.1523/JNEUROSCI.2515-11.2011

Copyright © 2011 the authors 0270-6474/11/3112716-11\$15.00/0

between the emergence of learning in individual neurons and in pursuit behavior. In the FEF_{SEM}, learning was episodic, emerging at different times in different neurons and sometimes vanishing by the end of a learning session. Even in the floccular complex, some Purkinje cells did not express learning until after the learned behavioral response had reached an asymptote. We conclude that the FEF_{SEM} and the floccular complex contain multiple learning processes with diverse time courses, with learning distributed temporally in both structures.

Materials and Methods

General methods. We report further analyses of two sets of data that have been published previously to investigate neural correlations of motor learning in pursuit eye movements (Medina and Lisberger, 2008, 2009; Li and Lisberger, 2011). Data were recorded from four adult male rhesus monkeys during a paradigm that induced directional learning in smooth pursuit eye movements. Two monkeys participated in recordings from the FEF_{SEM} and two in recordings from Purkinje cells in the floccular complex of the cerebellum. The experimental protocol for the two brain areas used the same general procedures and learning paradigm, with minor differences in technique that are mentioned at the relevant places in the following sections.

Throughout each daily experiment, the monkey sat comfortably in a primate chair with his head held stationary using implanted hardware. We monitored eye position using a magnetic scleral search coil system. The hardware for fixing the head, the eye coil, and the recording chamber was attached in sterile surgery with the monkey under isoflurane anesthesia (details described previously by Ramachandran and Lisberger, 2005). Postsurgical discomfort was mitigated by several days of treatment with opiates and nonsteroidal analgesics. The experimental procedures and protocols were approved in advance by the Institutional Animal Care and Use Committee of the University of California, San Francisco, and were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Experimental design. Experiments took place in a dimly lit room. Target presentation was different for the two brain areas, but the resulting learning was very similar. For recordings from the FEF_{SEM}, visual targets were presented on a Barco monitor (model number: CCID 7651 MkII) that was placed 40 cm in front of the eye and subtended $61 \times 42^\circ$ of visual field. For recordings from the floccular complex, visual targets were created by using an optical bench to image the light from a projector. The beam was deflected and moved using a mirror-galvanometer system, and was projected onto the back of a projection screen that was located 114 cm in front of the monkey and subtended $53 \times 53^\circ$ of visual field.

Visual stimuli were presented in a series of discrete trials. Each trial started when the monkey fixated a 0.5° white spot at the center of the screen for a randomized duration lasting between 500 and 1000 ms. Subsequently, the target underwent a step displacement followed by a ramp motion that lasted 750 ms. Finally, the target stopped at its final position for an additional 500 ms. Throughout each trial, monkeys were required to keep their gaze within a small window centered on the target. The size of the fixation window varied slightly between the recordings from the two structures (see Medina and Lisberger, 2008; Li and Lisberger, 2011). Failure to fulfill the eye position requirements resulted in immediate termination of the trial. At the end of each successfully completed trial, the animal received a small drop of water or juice.

Each learning experiment consisted of (1) characterization of the direction tuning of the neuron under study (~ 100 – 150 trials), (2) a baseline block to establish the detailed responses of the neuron before learning (~ 100 trials), and (3) a learning block (~ 250 – 300 trials). We describe the experimental design for recordings from the FEF_{SEM} first, and then list the minor differences for recordings from the floccular complex. The prelearning characterization block delivered 10 to 20 repetitions of step-ramp target motion at $20^\circ/\text{s}$ in each of the eight cardinal and oblique directions (Fig. 1A). During the subsequent baseline block, the target moved at $20^\circ/\text{s}$ along the cardinal axis closest to orthogonal to the preferred direction of the neuron under study. The example neuron that provided the data illustrated in Figure 1A preferred upward

pursuit, so the axis for the baseline block was horizontal. One direction along the baseline axis was designated as the probe direction (Fig. 1B, black arrow) and the other as the control direction (Fig. 1B, green arrow). In the baseline block, 55% of the baseline trials were “probe” trials and 45% were “control” trials. In the learning block, we introduced “learning” trials that provided an instructive stimulus to a learned change in the direction of pursuit (after Medina et al., 2005). In each learning trial, the target initially moved in the probe direction at $20^\circ/\text{s}$ for 250 ms. Then, the target underwent a change in direction caused by the addition of motion in the learning direction at $30^\circ/\text{s}$. The learning direction was chosen as the cardinal direction closest to the neuron’s preferred direction and was, by definition, orthogonal to the probe direction (Fig. 1B, bottom). Learning, probe, and control trials comprised 45%, 10%, and 45% of the total number of trials in the learning block for recordings from the FEF_{SEM}. In all blocks, the different types of trials were interleaved in random order.

There were a few minor differences in the experimental design for recordings from the floccular complex. First, the prelearning characterization block also served as the baseline block; as a result, only 12.5% of the trials in the baseline block featured target motion in the probe direction. Second, the learning block did not include target motion in the control direction, so that learning and probe trials comprised 90% and 10% of the trials. The schematic diagrams in Figure 1B summarize the differences in experimental design for the recordings from the two brain areas.

Data acquisition. Horizontal and vertical eye positions were sampled at 1 kHz and passed through analog differentiators with a cutoff of 25 Hz to yield the eye velocity traces. The eye velocity traces from each trial were examined on a computer screen to identify and excise any saccades, which were replaced by a linear segment of eye velocity from the values at the onset to the offset of the saccade.

Single-unit recordings from the FEF_{SEM} and the floccular complex were made, respectively, using quartz shielded tungsten electrodes from Thomas Scientific and glass shielded platinum-iridium microelectrodes manufactured in our laboratory. FEF_{SEM} neurons were characterized by robust, directionally tuned activity during smooth pursuit, and weak or nonexistent responses to saccades, eye position, or visual motion during fixation. Purkinje cells in the floccular complex were identified based on the modulation of their simple-spike firing during pursuit of sinusoidal target motion at 0.5 Hz and by the presence of complex spikes. The recorded voltage signals were amplified, filtered, and sorted offline into discrete units using methods described previously in detail (Medina and Lisberger, 2008; Li and Lisberger, 2011). For the analyses reported here, sorted waveforms were converted into spike trains with a temporal precision of 1 ms.

Calculation of neural and behavioral learning curves. Although all learning experiments produced statistically significant changes in mean eye velocity (significance level $p < 0.001$, Mann–Whitney U test), expression of neural learning was highly variable across the FEF_{SEM}, and to a lesser extent, the floccular complex. Thus, the preliminary step was to identify neurons with significant expressions of learning.

We used three different methods to identify neurons with significant expressions of learning. The three methods produced largely overlapping sets of neurons, and we included neurons that satisfied any of the three methods to make sure that we did not exclude neurons that expressed learning. All three methods compared neural responses from the complete set of probe trials in the baseline block with the first 100 learning trials from the learning block. First, we included neurons with a significant ($p < 0.01$, Mann–Whitney U test) change in mean spike count between the baseline block probe trials and the learning trials, measured in the interval from 100 to 320 ms after the onset of target motion. The first method identified 37/86 FEF_{SEM} neurons and 22/31 Purkinje cells. Second, we screened for neurons in which a significant change in activity was present only transiently during the block of learning trials. For each neuron, we determined whether learning caused an increase or decrease in the mean spike count (integrated across 100–320 ms), counted the number of learning trials whose spike count lay above the 95th percentile (for an increase in mean spike count) or below the fifth percentile (for a decrease) of the spike count distribution from the baseline block probe

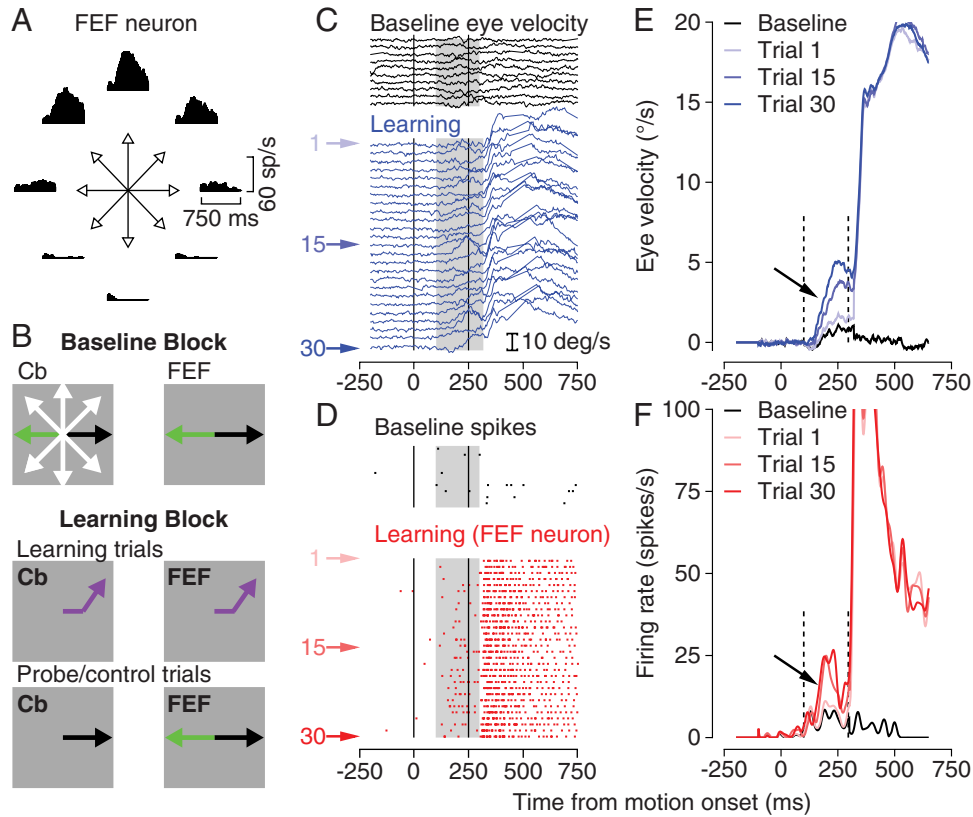


Figure 1. Examples of trial-over-trial changes in neural and behavioral responses during pursuit learning for a representative neuron in the FEF_{SEM}. **A**, The direction tuning of an example FEF_{SEM} neuron. Target motion began at the start of each histogram. sp/s, Spikes per second. **B**, Trial configurations for the baseline block and the learning block, shown separately for the FEF_{SEM} and floccular complex learning experiments. In this example experiment, the learning direction is upward, and the probe (black arrow) and control (green arrow) directions are, respectively, rightward and leftward. Cb, Cerebellum. **C**, Black and blue traces show eye velocity traces in the baseline block and the first 30 learning trials from the learning block. **D**, Black and red rasters show spike trains from the last 10 probe trials in the baseline block and the first 30 learning trials from the learning block. The order of trial occurrence runs from top to bottom. **E**, **F**, Blue and red traces show the smoothed eye velocity traces (**E**) and firing rate estimates (**F**) from the first, 15th, and 30th learning trials. The black traces plot the smoothed neural and behavioral data averaged across the last 10 probe trials from the baseline block. The diagonal arrows in **E** and **F** indicate the learned responses. The light shading in **C** and **D** and the vertical dashed lines in **E** and **F** indicate the analysis intervals.

trials, and included neurons with at least 11 such learning trials (1% chance occurrence). The second method added 9/86 FEF_{SEM} neurons and 2/31 Purkinje cells neurons for further study. Finally, to screen for neurons with learned responses that were brief in time during each individual trial, we smoothed the 0–320 ms portion of the spike train of each trial with a 50 ms rectangular filter, performed a millisecond-by-millisecond statistical comparison of the smoothed responses from the learning trials with those from the baseline block probe trials, and included all neurons with at least 50 time points where the two populations were significantly different (significance level $p < 0.01$, Mann–Whitney U test). The third method added 2/86 FEF_{SEM} neurons and no Purkinje cells. In total, 48/86 FEF_{SEM} neurons and 24/31 Purkinje cells satisfied at least one of the three criteria.

We used an adaptive algorithm that has been described previously in detail to generate a smoothed estimate of the underlying firing rate of each neuron during each pursuit trial as well as a smoothed estimate of how the activity changes across a learning session (Frank et al., 2002; Wirth et al., 2003). The adaptive algorithm estimates the firing rate iteratively based on a combination of the previous estimate and the presence or absence of a spike at each time point. Each update increases the likelihood that the estimated firing rate function would produce the observed spike train (Brown et al., 2001). The algorithm's virtues have been outlined by Frank et al. (2002), who demonstrated that it captures changes in neural responses over time accurately. By comparison with traditional smoothing methods such as convolution of the spike train with an exponential or Gaussian filter, the adaptive algorithm responds more quickly to trends in the data, is less sensitive to rapid fluctuations,

and seamlessly filters the data both within time in an individual trial and across consecutive trials.

In brief, the algorithm models the underlying firing rate of the neuron, $\lambda(t|\theta)$, on each trial as a cardinal spline with 39 control points, $\{\theta_{jj}\}_{j=1}^{39}$, located at 20 ms intervals, $\{t_{jj}\}_{j=1}^{39}$, from 100 ms before to 660 ms after the onset of target motion:

$$\lambda(t|\theta) = [v(t)^3 \ v(t)^2 \ v(t) \ 1] \times \begin{bmatrix} -0.5 & 1.5 & -1.5 & 0.5 \\ 1 & -2.5 & 2 & -0.5 \\ -0.5 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_{j-1} \\ \theta_j \\ \theta_{j+1} \\ \theta_{j+2} \end{bmatrix}. \quad (1)$$

The control points of the spline (θ_{j-1} through θ_{j+2}) used in Equation 1 place each time t within the interval between the times of θ_j and θ_{j+1} . We forced $\lambda(t|\theta)$ to be nonnegative to ensure a positive firing rate and we

$$\text{defined } v(t) = \frac{t - t_j}{t_{j+1} - t_j}.$$

The algorithm smoothes across time during a single trial by stepping through time and using the difference between the instantaneous observed number of spikes $dN(t)$ and the current state of the firing rate estimate to iteratively choose the spline that provides the best description of the data:

$$\theta_t = \theta_{t-1} + \varepsilon \frac{d\lambda}{d\theta} [dN(t) - \lambda(t|\theta_{t-1})\Delta t]. \quad (2)$$

Here, θ_t is the vector of control points for the analysis at time t , $dN(t)$ is 1 or 0 depending on the presence or absence of a spike, and $\lambda(t|\theta_{t-1})$ has units of spikes per second. After stepping through all time points in one trial, the final vector of control points specifies the estimate of underlying firing rate for that trial. The learning rate of the algorithm, ε , controls the extent to which the novel spiking information modifies the firing rate estimate. We chose ε to be 0.005, thereby allowing the firing rate estimate to be changed by at most 0.005 spikes per millisecond for each step of the algorithm. Our conclusions did not depend on the exact parameters chosen for the algorithm. Comparable results were produced using learning rates of 0.01 or 0.0025 and a 10 or 50 ms spacing between successive control points.

The algorithm smoothes across trials during a learning block by using the control points after analyzing the n th trial as a starting point for estimating the firing rate in the n th trial. For each set of data, we ran the algorithm forward by starting with the first trial in a block, filtering forward in time within that trial. The state of the control points after filtering a trial estimates the underlying firing rate for that trial and also serves as a starting point for filtering forward in time for the next trial, and so on. We then ran the algorithm backward by starting with the last time in the last trial of the block and filtering backward in time within that trial. We again estimated the underlying firing rate for a given trial as the state of the control points after filtering that trial, and used the control points from one trial as a starting point for filtering backward in time in the previous trial, and so on. The data block included all probe trials in the baseline block and all learning trials in the learning block. When the algorithm was run forward, we initialized the control points with the first 20 probe trials. When it was run backward in time, we initialized the control points from the mean firing rate across the last 20 learning trials. Running the algorithm forward and backward produced two separate firing rate estimates for each trial. These were averaged to yield the final firing rate estimate for that trial. Combining the forward and backward estimates prevented the smoothing algorithm from introducing any systematic time shifts into the final estimate of firing rate within individual trials or into the learning curves. Nonetheless, because the algorithm filters across trials, the estimate for Trial 1 includes information from future trials, and thus can reflect the subsequent learning.

To smooth the behavioral responses across a learning block, we convolved each millisecond of data across trials with an exponential filter. The decay constant of the filter was chosen to be -0.077 to match the parameters used for the adaptive algorithm. As with the neural data, the behavioral data were smoothed forward and backward across the learning block, and the results were averaged. We chose the exponential filter for the eye velocity data because it is more appropriate for continuous-valued data, and because the nature of the noise in eye velocity did not require the adaptive algorithm. We chose the adaptive algorithm for the neural data because it was derived for, and performed well, on the highly variable spike trains that arise from a point process. We also confirmed that our general conclusions were unchanged when the spike trains were smoothed with an exponential filter.

The learning curves quantified the magnitude of the neural or behavioral response as a function of the trial number. Each learning curve included data from the last nine probe trials in the baseline block, followed by data from the first 100 learning trials. To ensure that learning curves exclusively featured responses associated with learning, we focused the interval spanning 100–320 ms after the onset of target motion, because inherent delays within the pursuit circuit prevented the instructive change in target direction at 250 ms from directly affecting neural and behavioral responses until at least 70 ms later, at 320 ms. We measured the behavioral response on each trial as the distance traveled by the eye in the interval from 100 to 320 ms after the onset of target motion by calculating the integral of the eye velocity trace. For the neural responses, we estimated the smoothed number of spikes by integrating the firing rate from 100 to 300 ms after the onset of target motion. The shorter interval accounts for the ~ 20 ms resolution of the adaptive algorithm. Finally, we subtracted the average of the nine baseline values from each point along the learning curve and normalized the resulting curve so that all learning curves had a maximum value of 1.

Results

In recordings from both the FEF_{SEM} and the floccular complex, we searched for neurons that responded selectively to smooth pursuit and characterized responsive neurons according to their preferred direction during step-ramp target motion (Rashbass, 1961). We defined the preferred direction as the direction of target motion that elicited the largest mean firing rate, which was upward, or 90°, for the example neuron in Figure 1A.

We induced smooth pursuit learning by repeatedly presenting a target trajectory consisting of an initial 250 ms of motion at 20°/s in the probe direction, followed by the addition of a 30°/s velocity component in the learning direction for 500 ms. As described in the Materials and Methods section, we chose the learning direction to be the cardinal direction closest to the neuron's preferred direction and the probe direction to be orthogonal to the learning direction. For the neuron used to create Figure 1, target motion in the learning trials started in the rightward (probe) direction and then added a component of motion in the upward (learning) direction.

Our goal in the present paper differs from that in our prior reports, which documented the differences between mean eye velocity and neural firing after versus before directional pursuit learning (Medina and Lisberger, 2008, 2009; Li and Lisberger, 2011). Here, we treated learning as a dynamic process and evaluated how eye velocity and firing rate evolved across learning trials. Thus, most of our results are presented as a function of the number of the learning trials during the learning block. Because the monkey paced the experiment, absolute time was not under our control. We return to a conventional analysis of the mean learned firing rate and eye velocity as functions of time during single pursuit responses only in the final two figures of the paper.

Learned changes in eye velocity emerged gradually over the course of a learning session. In the example learning experiment in Figure 1, the vertical eye velocity traces from the last 10 probe trials in the baseline block were near zero (Fig. 1C, black traces) because target motion was exclusively to the right. During the first learning trial (Fig. 1C, top blue trace), the vertical eye velocity remained near zero until ~ 70 ms after the upward change in target direction. At that time, the vertical eye velocity exhibited a sharp upward deflection that we refer to as the visually guided eye movement because it was directly driven by the change in target direction. As we proceeded from first to the 30th learning trial, the general size and shape of the large visually guided eye movement were unaffected, but the eye velocity acquired an earlier upward deflection that peaked around the time of the change in target direction, 250 ms after the onset of target motion. This deflection in vertical eye velocity appeared too early in the trial to be part of the visually guided response; it also became progressively larger in learning trials that occurred later in the learning block. Therefore, the early upward eye velocity response is understood best as a learned response, rather than as the beginning of the visually guided response or as random eye velocity fluctuations.

Learned changes in firing rate also emerged gradually over the course of a learning session. Before learning, neurons showed little or no firing during baseline block probe trials that took the pursuit target and the eyes exclusively in a direction that was approximately orthogonal to the neuron's preferred direction (Fig. 1D, black raster). In the first learning trial (top row, red raster), the neuron retained a low level of activity up to 320 ms, 70 ms after the change in target direction. Starting 320 ms after the onset of target motion, the neuron fired vigorously because the change in target direction evoked a large visually guided eye

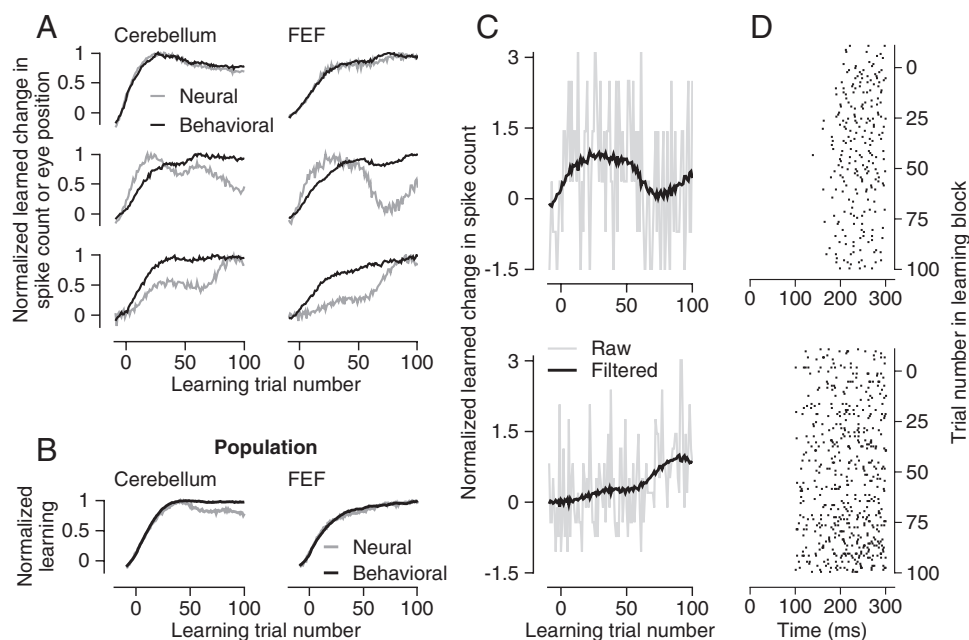


Figure 2. Comparison of neural and behavioral learning curves. Each graph includes black and gray traces that plot the normalized learned change in eye position or spike count as a function of the learning trial number. **A**, Comparison of neural and behavioral learning curves from six individual experiments. **B**, Comparison of normalized neural and behavioral learning curves averaged across the full set of experiments. **C**, Dark and light gray curves superimpose the smoothed learning curves and the spike counts from the original spike trains for the second and third example FEF_{SEM} neurons in **A**. **D**, Rasters showing the time of spikes during the analysis interval for the last 11 baseline trials and the first 100 learning trials for the neurons used in **C**. Time during each trial runs from left to right, and the trial sequence during the baseline and learning blocks runs from top to bottom.

movement in a direction close to its preferred direction. Across learning trials, the neural response during the visually guided eye movement did not change in size, but a small learned response developed in the interval between the onset of rightward target motion and the change in target direction (Fig. 1C, between the vertical black lines).

The data in Figure 1, *C* and *D*, are raw and unsmoothed. Figure 1, *E* and *F*, shows that the salient properties of the learned responses described above were evident after the data were smoothed across trials (see Materials and Methods). As the session proceeded from the first to the 30th learning trial, both the smoothed eye velocity and firing rate traces showed a gradual acquisition of learned responses (diagonal arrows) that preceded the onset of the instructive change in target direction. As a result of the forward and backward smoothing procedure across trials, a small learned response was already present in our smoothed representation of the neural and behavioral responses for the first learning trial.

Neural and behavioral learning curves

We portray the learning dynamics of individual neurons and pursuit behavior with learning curves, which measure the size of the normalized learned change in spike count or eye position as a function of the number of elapsed learning trials (see Materials and Methods). Figure 2A superimposes the neural (gray traces) and behavioral (black traces) learning curves for six different recording experiments, three each from the FEF_{SEM} and the floccular complex. For some neurons in both areas, the neural and behavioral learning curves followed almost identical time courses. But for many other neurons, the neural and behavioral learning curves diverged considerably. Across experiments, the shapes of the behavioral learning curves were fairly homogeneous and consisted of a rapid increase in the size of the learned component of the pursuit eye movement that occurred over approximately 20 learning trials and was maintained over the subsequent 80 trials. In contrast, the shapes of the neural learning curves

were much more heterogeneous. The divergence of neural and behavioral learning curves suggests that the learning dynamics of individual neurons in the FEF_{SEM} and the floccular complex reflect the dynamics of local learning processes, rather than merely changes in the animal's motor performance throughout the learning session.

Comparison of selected learning curves with the raw spike counts (Fig. 2C) and the original spike trains for the analysis interval (Fig. 2D) shows that the smoothed curves are a fair representation of the slower dynamics present in the unsmoothed data. For the top graphs in Figure 2, *C* and *D*, the early increase in the learning curve in Figure 2C is caused by the higher density of spikes in the part of the raster that surrounds Trial 25 in Figure 2D. The steady increase in the learning curve in the bottom panel of Figure 2C comes from the steady increase in the density of spikes moving from Trial 25 to Trial 100 in the bottom panel of Figure 2D. In addition to revealing the trends in the data in the face of considerable variation, we think that the smoothed curves provide two important indices of brain function. First, the adaptive filtering method estimates the underlying firing rate of a neuron with a highly variable spike train and provides a means to analyze changes in that underlying firing rate in association with learning. Second, the smoothed firing rate estimates the pooled output of a group of neurons with similar response properties and learning dynamics, thereby reflecting the signal seen by downstream areas.

Although the learning curves of individual neurons sometimes bore little resemblance to the behavioral learning curve, it was possible to recover the behavioral learning dynamics by averaging across neurons. Figure 2B reveals a high degree of similarity between the mean behavioral learning curve (black traces) and the average of the full set of learning curves from either the FEF_{SEM} or the floccular complex (gray traces). The average population learning curve from the floccular complex appears to fit the behavioral learning curve slightly less well than the average

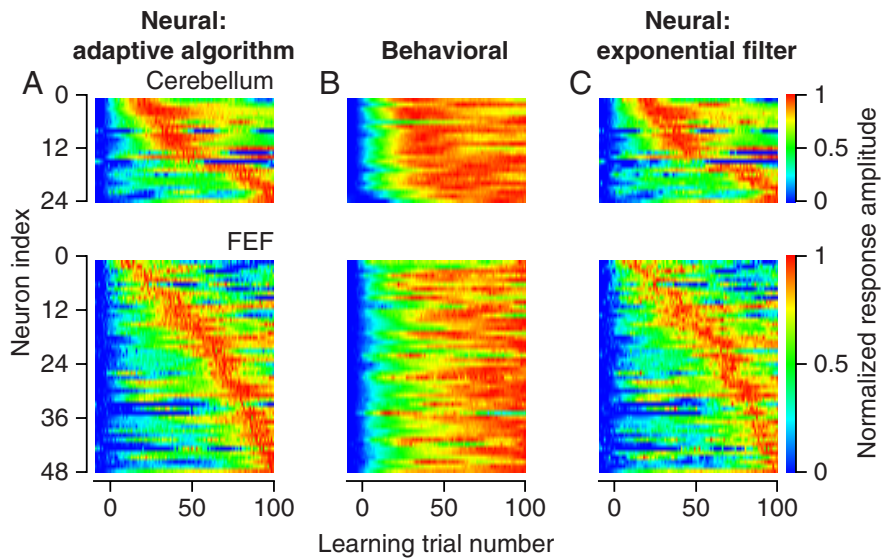


Figure 3. Neural and behavioral learning curves for all recordings from FEF_{SEM} and floccular complex. Each image uses color to show the value of the normalized learning curve and plots each learning curve as a horizontal line in the image. **A**, Neural learning curves based on the adaptive filtering algorithm. **B**, Behavioral learning curves. **C**, Neural learning curves obtained after smoothing the spike trains with a traditional exponential filter. Neural learning curves are ordered according to the trial when they first reached 95% of their peak value, and the behavioral learning curves are ranked according to the order of the neural learning curves. The top images show data from the floccular complex, and the bottom images show data from the FEF_{SEM}.

FEF_{SEM} learning curve, probably because there were half as many Purkinje cells as FEF_{SEM} neurons in our dataset. In support of this possibility, population learning curves generated from randomly bootstrapped samples of 24 FEF_{SEM} neurons accounted for, on average, 93.3% of the variance in the mean behavioral learning curve, similar to the 92.8% accounted for by the average learning curve from the Purkinje cells.

Comparison of learning dynamics in the FEF_{SEM} and the floccular complex

To visualize the dynamics of the full set of neural learning curves from the FEF_{SEM} and the floccular complex, we created the representation shown in Figure 3A. After ordering all of the learning curves from each brain area according to the trial when the learning curve first reached 95% of its maximal value, we plotted each neuron's learning curve as a horizontal line and represented the values of the learning curve using a color scale. For both the floccular complex and the FEF_{SEM}, the colored summary of the learning curves contains a diagonal red swath from top left to bottom right of the image. The learning curves in Figure 3A were obtained using the adaptive filter algorithm for the spike trains. The same features appear in the learning curves obtained after filtering the spike trains with an exponential filter (Fig. 3C).

The red swaths in Figure 3A indicate that (1) different neurons reach their peak learned responses after a different number of learning trials, (2) each neuron shows a learned response that peaks during a specific stretch of learning trials within the learning session, and (3) neurons with a peak early in the learning session frequently showed a decline in the expression of learning late in the learning session. The red swath of peaks is continuous for the FEF_{SEM}. The even tiling of the learning session by FEF_{SEM} neurons explains why the learned population response can grow smoothly across the learning session even though the expression of learning in individual neurons could be quite transient. The image for the floccular complex has a diagonal red swath that is slightly broken, giving the impression of two groups of Purkinje

cells that reach their peak learned responses early versus late in the learning block. In contrast to the diversity in the trajectories of the neural learning curves in both the floccular complex and the FEF_{SEM}, the behavioral learning curves (Fig. 3B) had fairly stereotyped trajectories and maintained near-peak values throughout much of the learning block. For some experiments, the neural expression of learning was weak at times during the learning experiment when the learned behavior was maximal. The opposite situation also occurred, although less frequently.

Two factors underlie the progressive shift in the peak responses across the population throughout the learning session, as documented in Figure 3A. First, neurons did not acquire their learned responses concurrently: some neurons reached their peak learned response after just a few learning trials, whereas others did not become highly active until near the end of the learning session. Second, after reaching a peak, the magnitude of the learned neural response

tended to decrease with additional training to an extent that varied across individual neurons.

To compare and contrast the rates at which the two brain areas acquired learning, we defined the “trial of learning acquisition” for neural and behavioral learning curves as the trial on which the learning curve first reached 75% of maximum. We then plotted the trial of acquisition for neural learning as a function of that for behavioral learning (Fig. 4A). The distance of each point from the oblique dashed line indicates the size of the discrepancy between the time courses of acquisition for neural and behavioral learning in units of number of trials. The location of each point below or above the unity line indicates whether that neuron reached 75% of its maximum learned response more or less quickly than did the behavior.

The FEF_{SEM} and the floccular complex differed in the extent to which the trial of neural learning acquisition matched that of the behavior. Neurons from the FEF_{SEM} (Fig. 4A, open symbols) were distributed fairly evenly throughout the plot, at all distances from the dashed unity line. They plotted approximately equally above or below the line, indicating that learning in FEF_{SEM} neurons could appear before or after behavioral learning. Most of the Purkinje cells from the floccular complex (filled symbols) plotted near the unity line, with a minority displaced quite far above the line, indicating that they started to express learning well after the acquisition of behavioral learning. We quantified the distributions from the FEF_{SEM} and floccular complex in Figure 4A by computing the “learning acquisition difference” as the difference between the trials on which the neural and behavioral learning curves first reached 75% of their maxima. In a histogram of the learning acquisition difference (Fig. 4A, diagonal marginal histogram), the data from the floccular complex were bimodal: 58% of the neurons acquired learning within 10 trials of the behavior, whereas another 21% lagged the behavior by >50 trials. The distribution for the FEF_{SEM} was flatter and broader, with approximately equal proportions of neurons that led or lagged the behavior. There were groups of FEF_{SEM} neurons with learning

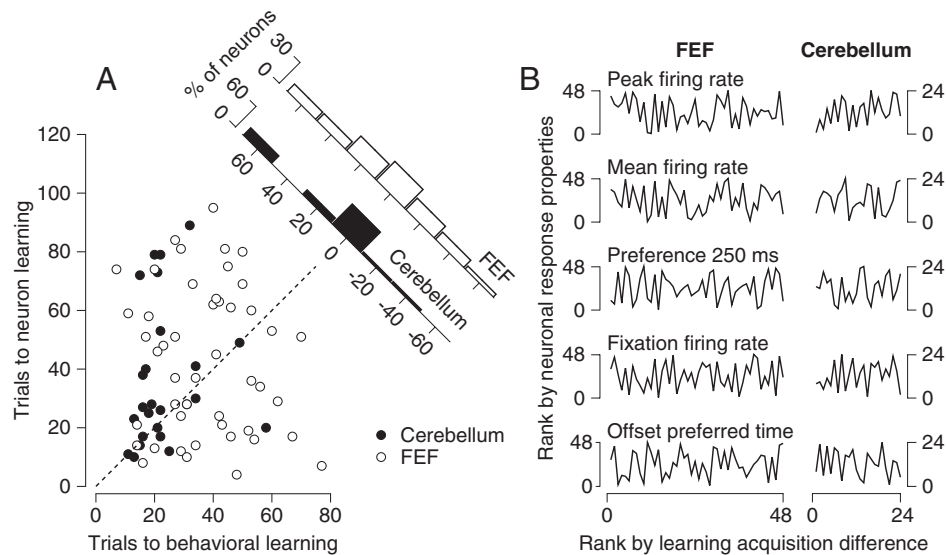


Figure 4. Quantitative comparison of the time course of acquisition of neural and behavioral learning. **A**, Each symbol shows data from an individual neuron and plots the trial of learning acquisition for the neural versus behavioral responses. Filled and open symbols show data from the floccular complex and FEF_{SEM}, respectively. The dashed line has a slope of 1. The oblique marginal histogram shows the distribution of the difference, in number of trials, between the acquisition of neural and behavioral learning. Note that these histograms summarize the distance of each neuron's position in the graph from the line of slope 1. **B**, Rank-order plots showing the relationship between each neuron's rank in the population in terms of various neural response properties measured before learning and the difference, in number of trials, between the acquisition of neural and behavioral learning. Peak firing rate: the peak mean firing rate of the neuron during pursuit trials in the prelearning block where the target moved in the learning direction. Mean firing rate: the difference in mean firing rate between prelearning pursuit in the probe direction versus in the learning direction, measured in the interval from 100 to 320 ms after the onset of target motion. Preference 250 ms: the mean firing rate of the neuron 250 ms after the onset of target motion during prelearning pursuit in the learning direction, normalized for the peak mean firing rate. Fixation firing rate: the mean firing rate measured during the last 200 ms of the fixation period before the onset of target motion. Offset preferred time: the absolute difference between the time after the onset of target motion at which the peak mean firing rate was reached and 250 ms.

acquisition differences that ranged from <10 (25% of the population) to >50 trials (12%), but no group was clearly predominant. Therefore, neurons in the floccular complex either acquired learning at approximately the same time as the behavior or at a considerably later time. The learning acquisition difference in FEF_{SEM} neurons was distributed more widely, with neurons leading or lagging the behavior by anywhere between a small to a large number of trials.

We obtained similar results when we repeated the analysis using thresholds of 65% or 85% of the maximum responses to determine the trial of neural and behavioral learning acquisition. We also obtained similar results when we estimated the learning acquisition difference by determining the number of trials one curve needed to be shifted to maximize the correlation between each pair of neural and behavioral learning curves. Finally, to control for the earlier acquisition of behavioral learning for the learning experiments in the floccular complex, we confirmed that our conclusions did not change when we performed the same analysis exclusively on experiments where the behavioral learning curve reached 75% of maximum within the first 40 learning trials.

We found very little evidence of a relationship between the learning acquisition difference of individual neurons and their other response properties. To create the presentation in Figure 4B, we ordered the members of each population of neurons according to their learning acquisition difference and five other response properties, used the orders to assign ranks for each property, and then plotted the rank according to each response property versus the rank for the learning acquisition difference. Smaller ranks for the learning acquisition difference were assigned to neurons that acquired learning faster. For the FEF_{SEM} (Fig. 4B, left column), we did not find any response properties that were related to the rank order for the learning acquisition difference. For the floccular complex (Fig. 4B, right column),

there was a positive relationship ($r = 0.40$, $p = 0.05$, Spearman's rank-order correlation) between the learning acquisition difference and the peak mean firing rate of the neuron during pursuit trials in the prelearning block where the target moved in the learning direction. There was no significant relationship with any of the other response properties we examined.

Inspection of Figure 3A suggests that neurons from the floccular complex whose responses peaked early in the learning session exhibited only a small decrease in activity as learning progressed, whereas many neurons in the FEF_{SEM} tended to undergo a larger and more rapid decrease in activity during later learning trials. To quantify the differences in the maintenance of learned responses between the FEF_{SEM} and the floccular complex, we plotted each neuron's activity across the first 30 learning trials against its activity across the last 30 learning trials (Fig. 5). In both brain areas, neurons with smaller learned responses early in the learning session tended to have larger learned responses by the end of the learning session. In the FEF_{SEM} (Fig. 5A), neurons showed decrements in their learned responses across the learning session if they had large learned response in the first 30 learning trials. As a result, there was a significant negative correlation of -0.44 ($p = 0.002$) between the sizes of the early and late learned responses across the FEF_{SEM} population. In contrast, neurons in the floccular complex maintained large learned responses until the end of the learning session, even if the learned response also was quite large early in the learning session (Fig. 5B). The correlation for the floccular complex was not significantly different from zero ($r = 0.27$, $p = 0.21$), and decreased to 0.02 if we removed one outlier data point located at $(-0.05, -0.42)$. The findings in Figure 5 did not depend on the exact interval chosen for early and late learning; using the first/last 20 or 40 learning trials led to similar conclusions.

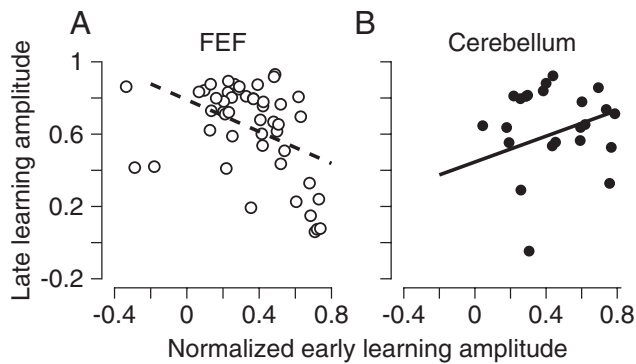


Figure 5. Relationship between the amplitude of neural learning early versus late in the learning session. Each symbol shows data for one neuron and plots the mean values of the normalized neural learning curve across the last 30 versus the first 30 learning trials. **A**, FEF_{SEM}; **B**, Floccular complex. The lines are the linear regression fits to the population data.

Time course of mean learned neural responses from the learning block

We now shift to a comparison of how the mean learned responses in the FEF_{SEM} and the floccular complex change on a millisecond time scale within an individual trial. Neurons within the two structures differ in their sensitivity to elapsed time relative to the onset of a pursuit movement evoked by standard step-ramp stimuli. Neurons in the FEF_{SEM} are most active at particular moments during pursuit, with different neurons encoding different times (Schoppik et al., 2008; Li and Lisberger, 2011). Purkinje cells in the floccular complex provide a time invariant representation of the velocity and acceleration of the eye movement (Shidara et al., 1993; Krauzlis and Lisberger, 1994).

In agreement with the earlier reports, we find a stronger time dependence in the learned responses of FEF_{SEM} neurons than in floccular Purkinje cells. The learned firing rate of the example neuron from the FEF_{SEM} (Fig. 6A, bottom) reached a peak before the time of the instructive change in target direction and then decreased, even though the learned eye velocity continued to increase until the time of the instructive change in target direction (Fig. 6B, bottom). In contrast, the learned response of the representative Purkinje cell from the floccular complex (Fig. 6A, top) increased steadily over almost 200 ms up to the time of the instructive change in target direction, in parallel with the learned change in eye velocity (Fig. 6B, top).

Across the population of neurons we recorded in the FEF_{SEM}, the learned responses of different neurons peaked at different times throughout the learned eye movement, whereas the peak times of the learned responses in the floccular complex were more uniform. In Figure 6C, we use color to indicate the normalized, smoothed learned firing rate responses of all members of the two populations of neurons. Each neuron's response is plotted as a horizontal line and the lines are ordered from top to bottom according to the millisecond during the learned neural response when their responses reached 95% of maximum. In the floccular complex, the red swath indicating the peaks of the learned neural responses lies between 250 and 296 ms after the onset of target motion, near the time of the instructive change in target direction. In the FEF_{SEM}, the red swath is narrower and more strongly diagonal, extending from 100 to 296 ms across the population of neurons. Thus, even during the brief learned eye movement, different FEF_{SEM} neurons were most active at distinct, transient moments. Collectively, the time of the peak responses in the FEF_{SEM} population spanned the duration of the learned eye movement. In contrast to the learned neural responses, ordering

the learned eye velocity traces according to the same ranking as the learned firing rates (Fig. 6D) failed to reveal a diagonal red swath. Therefore, the data provide no evidence that the temporal diversity in the peaks of the learned responses of FEF_{SEM} neurons resulted from heterogeneity in the peak times of the learned behavioral responses.

We suspect that intersubject variation accounts for the disparities in the behavioral responses in the data for the floccular complex versus the FEF_{SEM}. The experiments in the two structures were performed on different monkeys. Medina et al. (2005) showed that the time course of the mean learned eye velocity can differ across subjects. We do not think that differences in the behavior invalidate our neural findings because we found similar trends in the timing of the mean learned neural responses after we restricted our analysis to experiments where the mean learned eye velocity reached 95% of maximum after 280 ms.

Finally, we looked for a relationship between the peak time of the learned response in each FEF_{SEM} neuron and the time when the neuron was most active during the first 320 ms of prelearning pursuit in the direction that would become the learning direction. Plotting the time to 95% of the maximal response during prelearning pursuit as a function of that for the learned response revealed that almost all neurons plotted on or below the unity line (Fig. 7A). Thus, the learned responses of FEF_{SEM} neurons peak at the same time or later than their responses during prelearning pursuit. This intermediate result may represent a compromise between the intrinsic preferred timing of each individual neuron and the timing of the learned behavior caused by an instructive change in target direction at 250 ms.

Purkinje cells in the floccular complex tended to plot farther to the right than did neurons in the FEF_{SEM} (Fig. 7A,B), indicating that the temporal contingencies of the learned behavior had a stronger influence on the peak time of the learned responses in the floccular complex than in the FEF_{SEM}. To verify that the later peaks of the learned responses in the floccular complex were not related to the tendency toward a longer latency to peak for the mean learned eye velocity responses measured for the floccular complex experiments than for the FEF_{SEM} experiments (Fig. 6D), we focused on a subset of neurons with comparable latencies for behavioral learning. For experiments in which the learned eye velocity reached 95% of its maximum after 280 ms (Fig. 7, filled symbols), there still was a closer link between the timing of the learned behavior and the timing of learned responses in the floccular complex than in the FEF_{SEM}.

Discussion

Traditionally, behavioral learning has been thought of as a relatively simple process that is driven by smooth, continuous, and monotonic changes at a particular neural locus. Recent findings, however, are beginning to paint a more complex picture. Behavioral experiments in a number of movement systems have led to the conclusion that the smooth, continuous progression of behavioral learning arises from the cooperation of multiple components that operate on different time scales (Boyden et al., 2004; Smith et al., 2006; Lee and Schweighofer, 2009; Yang and Lisberger, 2010). In addition, cellular and molecular studies have uncovered an ever-increasing number of sites within different brain areas that are sensitive to various forms of physiological and structural plasticity with a wide range of temporal properties (Aizenman and Linden, 2000; Hansel and Linden, 2000; Zhang and Linden, 2006; Weeks et al., 2007; Pugh and Raman, 2008; Connor et al., 2009; Feldman, 2009). The neural correlates of the different learning components proposed by the behavioral studies have

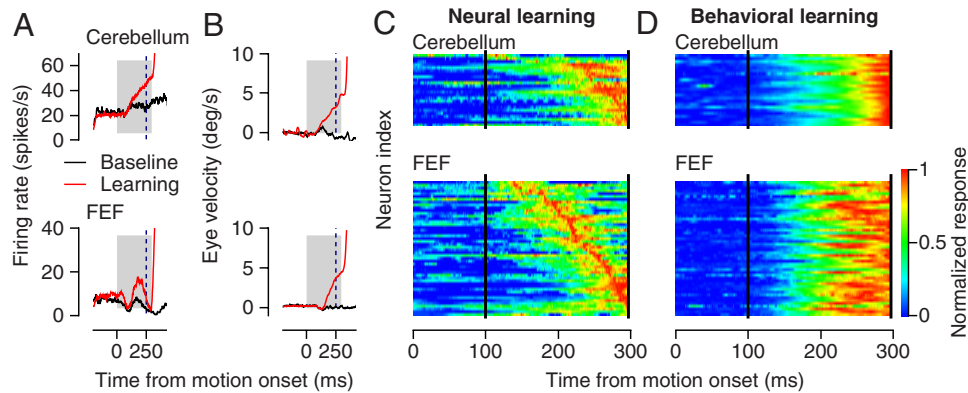


Figure 6. Millisecond time course of the mean learned neural and behavioral responses. **A**, Mean firing rates of a representative neuron recorded in the floccular complex (top) and FEF_{SEM} (bottom), plotted as a function of time from the onset of target motion. **B**, The mean eye velocities traces from the learning experiments shown in **A**. In **A** and **B**, black and red traces show data from probe trials in the baseline block and from learning trials. The gray shaded boxes delimit the interval plotted in **C** and **D**. **C, D**, Each image uses color to show normalized learned firing rate (**C**) or learned eye velocity (**D**). Each neuron's learned response occupies a horizontal line in the images. Learned neural responses are ordered according to the millisecond when they first reached 95% of their peak value, and the behavioral responses are ranked according to the order of the neural responses. The two black vertical lines at 100 and 296 ms in each of the panels of **C** and **D** are meant to delineate the analysis interval for the learned changes in firing rate and eye velocity. Because we smoothed the raw neural responses spanning 0–320 ms with a 50 ms rectangular filter, we had to discard the last 24 ms of the smoothed response, leaving a response that ended at 296 ms rather than at 320 ms. The top two images show data from the floccular complex, and the bottom two images show data from the FEF_{SEM}.

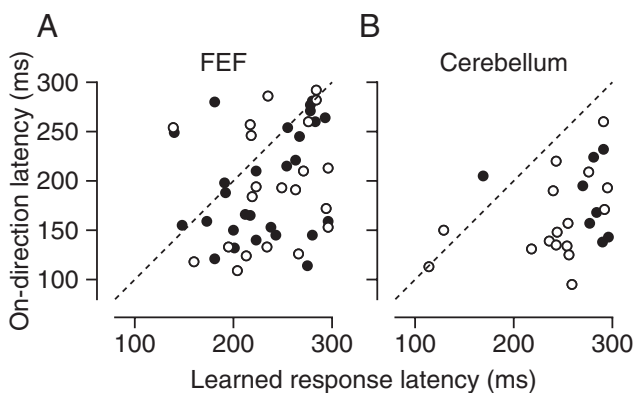


Figure 7. **A, B**, Millisecond time courses of prelearning and learned pursuit responses in the FEF_{SEM} (**A**) and the floccular complex (**B**). Each point shows data for one neuron and plots the time to 95% of the peak firing rate during prelearning pursuit in the learning direction versus the time to 95% of peak for the mean learned response. Filled symbols indicate the subset of experiments where the mean learned eye velocity first reached 95% of its maximal value after 280 ms. The dashed line has a slope of 1.

remained controversial, with some studies in favor of separate components being supported by different brain areas (Squire, 1992; Kassardjian et al., 2005; Paz et al., 2005), and others providing evidence that different neurons within a single anatomical structure could support separate components of learning (Chen and Wise, 1995; Wirth et al., 2003). Support for the latter possibility comes from our comparison of the time courses of neural learning in the FEF_{SEM} and the cerebellar flocculus in relation to learning in smooth pursuit eye movements.

We found that the learning expressed in individual neurons emerges over a time course that is heterogeneous in both the FEF_{SEM} and the floccular complex, and can be quite different from the smooth, continuous, and relatively quick time course of behavioral learning in pursuit eye movements. In both the FEF_{SEM} and the floccular complex, the progression of learned changes in neural responses over a learning session can be characterized best as “episodic.” As behavioral learning progresses continuously, expressions of learning appear (and sometimes disappear) in different neurons at different times. Some neurons

reach the peak of their learned response after just a few learning trials, whereas others do not acquire large learned responses until near the end of the learning session. In the FEF_{SEM}, the expressions of learning are especially transient. In the floccular complex, they are maintained more reliably through the learning session, but they still defy the standard expectation that learning in Purkinje cell firing will be acquired in lock step with the learned eye movement. One important implication of the episodic nature of neural learning is that the traditional comparison of neural responses before versus after learning has limited explanatory power and in some cases may be misleading.

The lack of a temporal correlation between behavioral and neural learning curves of individual neurons provides strong evidence for the prior suggestions that neural correlates of learning in the floccular complex and the FEF_{SEM} are not simply corollaries of the learned change in smooth eye movement (Medina and Lisberger, 2009; Li and Lisberger, 2011). Because the representations of learning in individual neurons do not simply track the eye movement response across trials, changes in neural firing must reflect the specific learning mechanisms and/or the signals present at the inputs to each neuron. Nevertheless, the population activity in both the FEF_{SEM} and the floccular complex follows the behavioral learning curve well, indicating that it still is reasonable to think that both of these areas play roles in creating the learned behavior.

Roles of the FEF_{SEM} and the floccular complex in pursuit learning

For both the FEF_{SEM} and the floccular complex, the temporal distribution of neural learning across a learning session provides insight into possible roles of these structures in pursuit learning. In the FEF_{SEM}, learning is expressed in a pattern that marches systematically across the relevant population of neurons as the learning session progresses. Because of its continuous tiling of the learning session, the FEF_{SEM} may be well suited for encoding the progression of behavioral learning. Prior work has suggested that the FEF_{SEM} plays a role in controlling the gain of visual-motor transmission during learning (Tanaka and Lisberger, 2001), and that learned changes in pursuit eye movement are closely related to the modulation of visual-motor gain (Chou and

Lisberger, 2004). One interpretation of our data could be that select groups of FEF_{SEM} neurons are involved in driving learning through changes in the gain of visual–motor transmission during different stages of the learning process.

The distribution of maximum learned responses in the FEF_{SEM} across the many minutes of a learning block fits with our earlier finding, on the much shorter time scale of hundreds of milliseconds, that different neurons show peak responses at different times during a single pursuit response (Schoppik et al., 2008; Li and Lisberger, 2011). Even in the first 320 ms of the learned response, we see the same tiling of time in the firing rates of neurons in the FEF_{SEM}. Different neurons reach peak learned responses at specific times that seem to be a compromise between the intrinsic timing properties of the individual neuron and the learned timing in the eye movement imposed by the change in target direction at 250 ms. We conclude that the population of neurons in the FEF_{SEM} represents time on multiple scales, both within the few hundred milliseconds of a learned pursuit response and across the many minutes of a pursuit learning session. Although we cannot be certain that changes in the firing patterns of neurons in the FEF_{SEM} drive learned changes in the behavior, we suspect that the FEF_{SEM} does play a causal role, mainly because of the short-latency effects of stimulation in the FEF_{SEM} on smooth eye movement (Gottlieb et al., 1993; Tanaka and Lisberger, 2001).

The discharge of Purkinje cells during pursuit is tightly linked with the velocity and acceleration of the eye movement (Shidara et al., 1993; Krauzlis and Lisberger, 1994), and the disinaptic connection from Purkinje cells to extraocular motoneurons (Highstein, 1973) implies that learned changes in floccular output will cause learned changes in eye velocity. Therefore, we had expected floccular learning curves to be quite different from those in the FEF_{SEM} and to follow the behavioral learning curves closely. In line with our expectations, many Purkinje cells acquired learned responses early in the learning block and retained large learned responses through the learning session. We also found a tight link between the trajectories of Purkinje cell firing rates and the learned eye velocity during the first 320 ms of learning trials. Nevertheless, the floccular complex also showed the episodic pattern of learning seen in the FEF_{SEM} to some degree, because a small group of Purkinje cells acquired learned responses after behavioral learning had reached an asymptote. Perhaps neural learning dynamics become less episodic and more like the behavior as signals move closer to the output of the motor circuit. For example, neurons in the deep cerebellar nuclei express learning that progresses in step with the oculomotor behavior during saccade adaptation (Inaba et al., 2003; Scudder and McGee, 2003), as do neurons in the nucleus reticularis tegmenti pontis (Takeichi et al., 2005). However, differences in the representations of learning in the FEF_{SEM} and the floccular complex could also be attributed to their involvement in anatomically distinct pathways within the pursuit circuit (Lynch and Tian, 2006; Ono and Mustari, 2009; for review, see Thier and Möck, 2006).

Learning as a multirate process

We have demonstrated the existence of multiple components of neural learning that operate on different time scales and have shown how different times throughout the learning process are represented across neural populations in pursuit-related areas of the motor cortex and the cerebellum. However, our data go beyond simply supporting the existence of multiple components. From a behavioral perspective, our identification of neural pro-

cesses with diverse learning dynamics may have important implications for the formation, organization, and storage of motor memories. Memories are acquired and maintained over a remarkably broad range of timescales. Part or all of the memory can be forgotten within seconds or can persist for days or even years. Recordings of the time course of neural learning provide an opportunity to test the idea that different components of memory are acquired and stored in different neural structures.

In motor learning, the cerebellar cortex and deep cerebellar nuclei are thought to support learning on short and longer time scales, respectively (Miles and Lisberger, 1981; Krupa et al., 1993; Raymond et al., 1996; Medina et al., 2001; Nagao and Kitazawa, 2003; Kassardjian et al., 2005). The motor cortex is believed to maintain long-term motor memories (Wise et al., 1998; Muellbacher et al., 2002; Paz et al., 2005; Richardson et al., 2006; Hadipour-Niktarash et al., 2007; Xu et al., 2009). Our data contradict the idea that each brain area is involved exclusively in a particular stage of memory formation. For example, the late acquisition of learned changes in some floccular Purkinje cells raises the possibility that long-term memories are stored not only in the deep cerebellar nuclei, but also in the cerebellar cortex. The early acquisition of learned changes in some FEF_{SEM} neurons suggests that the motor cortex may be involved in the formation of short-term memories, in addition to its established role in long-term memory storage.

Of course, the dense interconnectivity of the pursuit circuit allows expressions of learning in any given brain area to reflect inputs from other areas, rather than local plasticity. However, dense interconnectivity would be expected to cause homogeneity in the time of acquisition of learned neural responses, rather than the heterogeneity reported here. Therefore, we suggest that the diversity in the time course of learning across the neural populations arises from differences in the temporal properties of the synaptic plasticity mechanisms and the neural inputs to each individual cell.

References

- Aizenman CD, Linden DJ (2000) Rapid, synaptically driven increases in the intrinsic excitability of cerebellar deep nuclear neurons. *Nat Neurosci* 3:109–111.
- Boyden ES, Katoh A, Raymond JL (2004) Cerebellum-dependent learning: the role of multiple plasticity mechanisms. *Annu Rev Neurosci* 27:581–609.
- Brown EN, Nguyen DP, Frank LM, Wilson MA, Solo V (2001) An analysis of neural receptive field plasticity by point process adaptive filtering. *Proc Natl Acad Sci U S A* 98:12261–12266.
- Chen LL, Wise SP (1995) Supplementary eye field contrasted with the frontal eye field during acquisition of conditional oculomotor associations. *J Neurophysiol* 73:1122–1134.
- Chou IH, Lisberger SG (2004) The role of the frontal pursuit area in learning in smooth pursuit eye movements. *J Neurosci* 24:4124–4133.
- Connor S, Bloomfield J, LeBoutillier JC, Thompson RF, Petit TL, Weeks AC (2009) Eyeblink conditioning leads to fewer synapses in the rabbit cerebellar cortex. *Behav Neurosci* 123:856–862.
- Doyon J, Benali H (2005) Reorganization and plasticity in the adult brain during learning of motor skills. *Curr Opin Neurobiol* 15:161–167.
- Feldman DE (2009) Synaptic mechanisms for plasticity in neocortex. *Annu Rev Neurosci* 32:33–55.
- Frank LM, Eden UT, Solo V, Wilson MA, Brown EN (2002) Contrasting patterns of receptive field plasticity in the hippocampus and the entorhinal cortex: an adaptive filtering approach. *J Neurosci* 22:3817–3830.
- Gottlieb JP, Bruce CJ, MacAvoy MG (1993) Smooth eye movements evoked by stimulation in the macaque frontal eye field. *J Neurophysiol* 69:786–799.
- Hadipour-Niktarash A, Lee CK, Desmond JE, Shadmehr R (2007) Impairment of retention but not acquisition of a visuomotor skill

- through time-dependent disruption of primary motor cortex. *J Neurosci* 49:13413–13419.
- Hansel C, Linden DJ (2000) Long-term depression of the cerebellar climbing fiber–Purkinje neuron synapse. *Neuron* 26:473–482.
- Highstein SM (1973) Synaptic linkage in the vestibulo-ocular and cerebello-vestibular pathways to the VIth nucleus in the rabbit. *Exp Brain Res* 17:301–314.
- Inaba N, Iwamoto Y, Yoshida K (2003) Changes in cerebellar fastigial burst activity related to saccadic gain adaptation in the monkey. *Neurosci Res* 46:359–368.
- Kahlon M, Lisberger SG (2000) Changes in the responses of Purkinje cells in the floccular complex of monkeys after motor learning in smooth pursuit eye movements. *J Neurophysiol* 84:2945–2960.
- Kassardjian CD, Tan YF, Chung JY, Heskin R, Peterson MJ, Broussard DM (2005) The site of a motor memory shifts with consolidation. *J Neurosci* 25:7979–7985.
- Keating EG (1991) Frontal eye field lesions impair predictive and visually guided pursuit eye movements. *Exp Brain Res* 86:311–323.
- Krauzlis RJ, Lisberger SG (1994) Simple spike responses of gaze velocity Purkinje cells in the floccular lobe of the monkey during the onset and offset of pursuit eye movements. *J Neurophysiol* 72:2045–2050.
- Krupa DJ, Thompson JK, Thompson RF (1993) Localization of a memory trace in the mammalian brain. *Science* 260:989–991.
- Lee JY, Schweighofer N (2009) Dual adaptation supports a parallel architecture of motor memory. *J Neurosci* 29:10396–10404.
- Li JX, Lisberger SG (2011) Learned timing of motor behavior in the smooth eye movement region of the frontal eye fields. *Neuron* 69:159–169.
- Lisberger SG (1994) Neural basis for motor learning in the vestibuloocular reflex of primates. III. Computational and behavioral analysis of the sites of learning. *J Neurophysiol* 72:974–998.
- Lynch JC, Tian JR (2006) Cortico-cortical networks and cortico-subcortical loops for the higher control of eye movements. *Prog Brain Res* 151:461–501.
- MacAvoy MG, Gottlieb JP, Bruce CJ (1991) Smooth-pursuit eye movement representation in the primate frontal eye field. *Cereb Cortex* 1:95–102.
- Medina JF, Lisberger SG (2008) Links from complex spikes to local plasticity and motor learning in the cerebellum of awake-behaving monkeys. *Nat Neurosci* 11:1185–1192.
- Medina JF, Lisberger SG (2009) Encoding and decoding of learned smooth-pursuit eye movements in the floccular complex of the monkey cerebellum. *J Neurosci* 102:2039–2054.
- Medina JF, Garcia KS, Mauk MD (2001) A mechanism for savings in the cerebellum. *J Neurosci* 21:4081–4089.
- Medina JF, Carey MR, Lisberger SG (2005) The representation of time for motor learning. *Neuron* 45:157–167.
- Miles FA, Lisberger SG (1981) Plasticity in the vestibulo-ocular reflex: a new hypothesis. *Annu Rev Neurosci* 4:273–299.
- Muellbacher W, Ziemann U, Wissel J, Dang N, Kofler M, Facchini S, Boroojerdi B, Poewe W, Hallett M (2002) Early consolidation in human primary motor cortex. *Nature* 415:640–644.
- Nagao S, Kitazawa H (2003) Effects of reversible shutdown of the monkey flocculus on the retention of adaptation of the horizontal vestibulo-ocular reflex. *Neuroscience* 118:563–570.
- Ono S, Mustari MJ (2009) Smooth pursuit-related information processing in frontal eye field neurons that project to the NRTP. *Cereb Cortex* 19:1186–1197.
- Paz R, Natan C, Boraud T, Bergman H, Vaadia E (2005) Emerging patterns of neuronal responses in supplementary and primary motor areas during sensorimotor adaptation. *J Neurosci* 25:10941–10951.
- Pugh JR, Raman IM (2008) Mechanisms of potentiation of mossy fiber EPSCs in the cerebellar nuclei by coincident synaptic excitation and inhibition. *J Neurosci* 28:10549–10560.
- Ramachandran R, Lisberger SG (2005) Normal performance and expression of learning in the vestibulo-ocular reflex (VOR) at high frequencies. *J Neurophysiol* 93:2028–2038.
- Rambold H, Churchland A, Selig Y, Jasmin L, Lisberger SG (2002) Partial ablations of the flocculus and ventral paraflocculus in monkeys cause linked deficits in smooth pursuit eye movements and adaptive modification of the VOR. *J Neurophysiol* 87:912–924.
- Rashbass C (1961) The relationship between saccadic and smooth tracking eye movements. *J Physiol* 159:326–338.
- Raymond JL, Lisberger SG, Mauk MD (1996) The cerebellum: a neuronal learning machine? *Science* 272:1126–1131.
- Richardson AG, Overduin SA, Valero-Cabré A, Padoa-Schioppa C, Pascual-Leone A, Bizzi E, Press DZ (2006) Disruption of primary motor cortex before learning impairs memory of movement dynamics. *J Neurosci* 26:12466–12470.
- Schoppik D, Nagel KI, Lisberger SG (2008) Cortical mechanisms of smooth eye movements revealed by dynamic covariations of neural and behavioral responses. *Neuron* 58:248–260.
- Scudder CA, McGee DM (2003) Adaptive modification of saccade size produces correlated changes in the discharges of fastigial nucleus neurons. *J Neurophysiol* 90:1011–1026.
- Shidara M, Kawano K, Gomi H, Kawato M (1993) Inverse-dynamics model eye movement control by Purkinje cells in the cerebellum. *Nature* 365:50–52.
- Smith MA, Ghazizadeh A, Shadmehr R (2006) Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biology* 4:e179.
- Squire LR (1992) Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol Rev* 99:195–231.
- Takeichi N, Kaneko CR, Fuchs AF (2005) Discharge of monkey nucleus reticularis pontis neurons changes during saccade adaptation. *J Neurophysiol* 94:1938–1951.
- Tanaka M, Lisberger SG (2001) Regulation of the gain of visually guided smooth-pursuit eye movements by frontal cortex. *Nature* 409:191–194.
- Tanaka M, Lisberger SG (2002) Enhancement of multiple components of pursuit eye movement by microstimulation in the arcuate frontal pursuit area in monkeys. *J Neurophysiol* 87:802–818.
- Thier P, Möck M (2006) The oculomotor role of the pontine nuclei and the nucleus reticularis tegmenti pontis. *Prog Brain Res* 151:293–320.
- Weeks AC, Connor S, Hinchcliff R, LeBoutillier JC, Thompson RF, Petit TL (2007) Eye-blink conditioning is associated with changes in synaptic ultrastructure in the rabbit interpositus nuclei. *Learn Mem* 14:385–389.
- Westheimer G, Blair SM (1973) Oculomotor defects in cerebellectomized monkeys. *Invest Ophthalmol* 12:618–621.
- Wirth S, Yanike M, Frank LM, Smith AC, Brown EN, Suzuki WA (2003) Single neurons in the monkey hippocampus and learning of new associations. *Science* 300:1578–1581.
- Wise SP, Moody SL, Blomstrom KJ, Mitz AR (1998) Changes in motor cortical activity during visuomotor adaptation. *Exp Brain Res* 121:285–299.
- Xu T, Yu X, Perlik AJ, Tobin WF, Zweig JA, Tennant K, Jones T, Zuo Y (2009) Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature* 462:915–919.
- Yang Y, Lisberger SG (2010) Learning on multiple time scales in smooth pursuit eye movements. *J Neurophysiol* 104:2850–2862.
- Zhang W, Linden DJ (2006) Long-term depression at the mossy fiber-deep cerebellar nucleus synapse. *J Neurosci* 26:6935–6944.