Neural Networks 47 (2013) 64–71

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

2013 Special Issue

Nucleo-olivary inhibition balances the interaction between the reactive and adaptive layers in motor control

Ivan Herreros [a], Paul F.M.J. Verschure [a,b,*]

[a] Laboratory of Synthetic Perceptive Emotive and Cognitive Systems (SPECS), Centre of Autonomous Systems and Neuro-Robotics (N-RAS), Department of Technology, Universitat Pompeu Fabra (UPF), Barcelona 08018, Spain

[b] Catalan Institute for Research and Advanced Studies (ICREA), Spain

ARTICLE INFO

Keywords:
Cerebellum
Classical conditioning
Nucleo-olivary inhibition
Motor control
Adaptive reflexes

ABSTRACT

In the acquisition of adaptive motor reflexes to aversive stimuli, the cerebellar output fulfills a double purpose: it controls a motor response and it relays a sensory prediction. However, the question of how these two apparently incompatible goals might be achieved by the same cerebellar area remains open. Here we propose a solution where the inhibition of the Inferior Olive (IO) by the cerebellar Deep Nuclei (DN) translates the motor command signal into a sensory prediction allowing a single cerebellar area to simultaneously tackle both aspects of the problem: execution and prediction. We demonstrate that having a graded error signal, the gain of the Nucleo-Olivary Inhibition (NOI) balances the generation of the response between the cerebellar and the reflexive controllers or, in other words, between the adaptive and the reactive layers of behavior. Moreover, we show that the resulting system is fully autonomous and can either acquire or erase adaptive responses according to their utility.

1. Introduction

The execution of an avoidance action seems to involve both sensory prediction and motor control: the prediction of a noxious stimulus triggers an anticipatory motor command. A similar division between sensory prediction and actuation is also found in control theory when a forward model provides predicted feedback to a feedback controller (Miall, Weir, Wolpert, & Stein, 1993). In the latter case, the reactive commands of the feedback controller are caused by the factual feedback anticipated by the forward model. On the contrary, in the case of an avoidance action, a common sense interpretation suggests that the predicted sensory event is *counterfactual*, i.e., not the factual sensory event predicted but the one that would be perceived without the avoidance action. Here we will show that to understand the role of the cerebellum in Avoidance Learning (AL) one might have to drop this assumption.

Acquisition of anticipatory responses has been extensively studied with the paradigm of Pavlovian classical conditioning (Pavlov & Anrep, 1927), e.g., classical conditioning of the eyeblink reflex (Gormezano, Prokasy, & Thompson, 1987) (henceforth, eyeblink conditioning). In eyeblink conditioning a neutral cue such as a sound or a light, the Conditioning Stimulus (CS), precedes by a fixed time-interval the delivery of a noxious Unconditioned Stimulus (US) to the eye, e.g., a peri-orbital electric shock. The US occurrence triggers a reflexive protective action (the closure of the eyelid) that constitutes the Unconditioned Response (UR). After a number of paired CS–US repetitions, the subject reacts to the delivery of the CS by closing the eyelids in anticipation of the expected US, i.e., producing a Conditioned Response (CR) (Gormezano et al., 1987; Mackintosh, 1974; Pavlov & Anrep, 1927). Once acquired, CRs can be deleted by extinction training, i.e., presenting CSs not followed by the US.

There is broad agreement that the substrate of learning in eyeblink conditioning is located in the cerebellum (Christian & Thompson, 2003; Yeo & Hesslow, 1998). The well known cerebellar circuitry (Eccles, Ito, & Szentágothai, 1967) helped to accurately delineate the neural pathways of CS, US and CR (Mauk, Steinmetz, & Thompson, 1986; Steinmetz, Lavond, & Thompson, 1985). The roles of the different stimuli accord with Marr–Albus–Ito cerebellar learning theory (Albus, 1971; Ito, Sakurai, & Tongroach, 1982; Marr, 1969): the US signal relayed by the IO reaches the cerebellar cortex through the climbing fibers where it induces plasticity at the synapses of the parallel fibers that transmit the CS information. After repeated coincidence of these two signals, the Purkinje cells – the sole output of the cerebellar cortex – acquire a response to the CS, namely, a drop in their firing activity, that drives the behavioral CR (Jirenhed, Bengtsson, & Hesslow, 2007). Moreover, as with the overt CR, extinction training suppresses the Purkinje cell response.

* Corresponding author at: Laboratory of Synthetic Perceptive Emotive and Cognitive Systems (SPECS), Centre of Autonomous Systems and Neuro-Robotics (N-RAS), Department of Technology, Universitat Pompeu Fabra (UPF), Barcelona 08018, Spain.

E-mail address: Paul.verschure@upf.edu (P.F.M.J. Verschure).

0893-6080/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2013.01.026

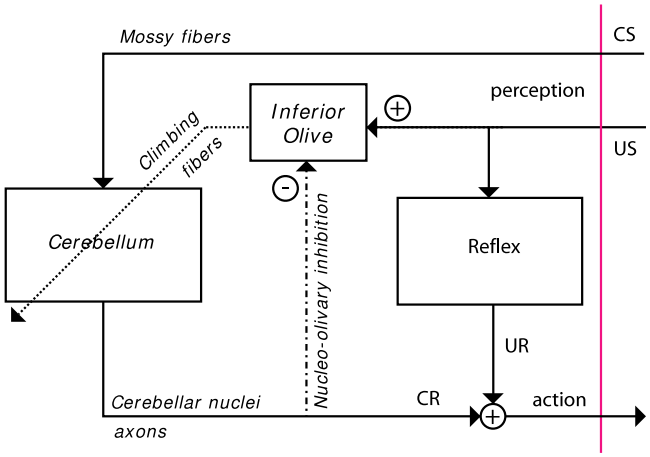

Fig. 1. Layout of the cerebellar controller plus reflexive arc. Information pathways are tagged according to the inputs outputs that they relay in classical conditioning. The labels in italic identify anatomically information pathways and processes. For a detailed description of the cerebellar architecture see Eccles et al. (1967).

Learning in classical conditioning regards sensory prediction. As the Rescorla–Wagner model formalized, animals learn in classical conditioning only when events violate their expectations (Rescorla & Wagner, 1972). Therefore, to support this kind of learning the cerebellum must acquire and generate sensory predictions. In general, according to the adaptive filter theory, cerebellar learning is explained in terms of de-correlation (Fujita, 1982). A corollary of this theory is that the cerebellum only learns when the IO activity is perturbed from baseline. In this context, the inhibitory connections from the cerebellar deep nuclear cells to the Inferior Olive, the Nucleo Olivary Inhibition (NOI) (Andersson, Garwicz, & Hesslow, 1988), are key to interpret cerebellar learning as the acquisition of sensory predictions. The NOI subtracts the cerebellar output relayed by the DN from the US signal reaching the IO, such that if both the signals match, they cancel each other leaving IO activity at baseline. Therefore, in eyeblink conditioning, if after the CS either the excitation produced by the US or the inhibition produced by the CR (via the NOI) outweighs the other, the perturbation of the IO activity would recruit cerebellar plasticity such that in following trials IO activity will remain closer to baseline. Remarkably, the NOI has an unusual long latency for a monosynaptic transmission in the order of the tens of milliseconds (Hesslow, 1986).

Regarding motor control, it is well-established that the output of the cerebellum drives the CR (Hesslow, 1994). In itself, this does not contradict the sensory prediction interpretation if the predicted US stimulus and the amplitude of the CR are correlated (although it is not obvious why such a correlation should exist). In other words, since correlation between neural activity and stimulus intensity – or action amplitude – is interpreted as evidence for the neural activity *coding* the stimulus – or the action – then, in classical conditioning, the cerebellar output may code both the perception *and* the response if perception and response are themselves correlated. However, the question remains whether the NOI, fundamental for sensory prediction, is functional from a motor learning perspective. AL, which, as a paradigm, is closely related to classical conditioning, serves us to address this issue.

In a classical conditioning preparation the CR is required to not ameliorate or reduce the noxiousness of the US. For instance, with a peri-orbital shock US, the CR has no effect in reducing the intensity of the shock. In AL, the CR modifies the effect of the US. For instance, if we use an airpuff without restraining the eyelids, then the effective or sensed intensity of the US will decrease as the CR increases the degree of eyelid closure, i.e., the noxiousness of the US will diminish as it reaches a more protected cornea. Therefore, whereas in classical conditioning the CR and the US can only be compared internally (and by means of the NOI), in AL an implicit *comparison* between action and stimulus takes place in the external world. This difference between classical conditioning and AL is not always explicitly made in the literature, since some eyeblink conditioning studies, specially with humans subjects, are made with an airpuff and an unconstrained CR (Clark & Squire, 1998).

However, attempting to apply the cerebellar microcircuit studied classical conditioning to a task of AL raises a series of questions.

First, if the cerebellum outputs a motor command and the IO receives a peripheral sensory signal, then the NOI performs a non consistent comparison between information from different domains. In such a case, why should the temporal profile of the signal masking a phasic US be similar to the motor command controlling the eyelid muscles? (Lepora, Porrill, Yeo, & Dean, 2010). Note that the same inconsistency of the temporal dynamics appears when we consider the avoidance of a noxious stimulus as a comparison performed in the external world. E.g., the temporal profile of the eyelid closure and the physical US stimulus might be different.

And second, AL introduces a contingency between the motor action and the sensory prediction: the CR diminishes the effective intensity of the US. We refer to this link as the behavioral negative feedback loop in contrast to the internal negative feedback provided by the NOI. But if the behavioral learning can avoid the US, what is the use of the internal negative feedback? Remark that in cases where avoidance can be complete (to hit against a wall or to completely avoid it) the role of the NOI is not evident, i.e., since both negative feedback loops are superposed, the NOI might halt learning before it leads to the total avoidance of the US.

However, it has been shown both with modeling and animal preparations that inactivation of the NOI prevents extinction in classical conditioning (Bengtsson & Hesslow, 2006; Medina, Nores, & Mauk, 2002). Extrapolating this result to AL, then the NOI has the functional role of suppressing acquired responses that are no longer adaptive. Therefore, even though it could be possible for a cerebellar microcircuit lacking the NOI to optimally acquire a response in AL, such circuit would require an extra-cerebellar brain structure to generate the signal driving extinction. In other words, in the absence of an external signal reflecting the cost of an unnecessary avoidance action, this signal, playing the role of a hypothetical 'negative US', has to be computed internally, and the NOI provides a means for its generation.

We propose that the key to reconcile sensory prediction and motor control lies in the nature of the US signal. Considering a graded rather than an all-or-none US signal, the NOI can halt learning once the US intensity drops below a certain *safety* level, that is, once the US is as mild as to lose its noxiousness. Moreover, this residual signal can play an important functional role, i.e., in a trial-by-trial basis, it can validate the suitability of the anticipatory action. For instance, in the case of AL of the eyeblink response, once the eyelids are closed, perceiving the airpuff confirms the suitability of keep triggering the anticipatory action the next time the CS is perceived.

To summarize, we propose that the NOI allows balancing the level of control between a reactive and an adaptive layer. We validate this proposal in a series of simulations where a robot has to perform a collision avoidance task in a track with a single turn. For the adaptive layer we use a controller based on the anatomy of the cerebellum (Fig. 1) (Eccles et al., 1967). Using the principles behind adaptive filter modeling of the cerebellum, we implement an analysis–synthesis filter with a de-correlation learning rule (Dean, Porrill, Ekerot, & Jörntell, 2010). With this setup, we study the effects of different parametrizations of the NOI gain, showing that it fixes the balance between reactive and adaptive actions, and that besides being required for extinction, the NOI is fundamental for the correct timing of the adaptive responses.

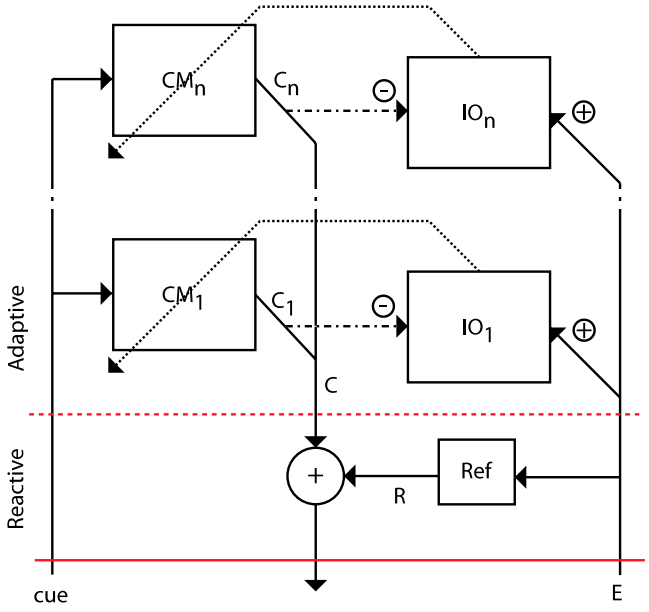

Fig. 2. Computational architecture with the reactive and adaptive controllers including N cerebellar microcircuits (CM), N IO components and a reflex arc (Ref). Each microcircuit generates an output command C_i. The individual commands are averaged into C and then added to R to generate the final action.

2. Methods

2.1. Cerebellar model

The model of the cerebellum consists of a set of parallel cerebellar microcircuits, each one connected to its IO component. Each cerebellar microcircuit encapsulates information processing from cerebellar cortex and cerebellar nuclei together. The inputs displayed in Fig. 2 correspond to the mossy fiber and the climbing fiber pathways, and relay the cue and the error signals, respectively. In nature, the output of a microcircuit module will be carried by the axons of the deep-nuclear cells (for a review of the cerebellar cytoarchitecture see Eccles et al., 1967).

We implemented each cerebellar microcircuit as an analysis–synthesis adaptive filter (see Dean et al., 2010 for a review), where the information coming through the mossy fiber pathway is mapped into deep-nuclear activity and adjusted according to the teaching signal provided by the climbing-fiber signal. Even though the assumptions inherent to this model are described in a series of publications (Dean et al., 2010; Fujita, 1982; Porrill & Dean, 2008) we briefly describe them here. The cerebellar cortex acts as a filter that maps mossy-fiber activity into the Purkinje cell output. Upon entering the cerebellar cortex, the mossy fiber information is expanded into multiple components or *basis* into the granular layer. Such basis arise by means of the interaction between the mossy fibers, the excitatory granule cells and the inhibitory Golgi cells (Medina & Mauk, 2000; Yamazaki & Tanaka, 2007). Different factors such as the diversity in the connectivity between these cells (Mapelli, Gandolfi, & D'Angelo, 2010) and in the synaptic gain distribution (Crowley, Fioravante, & Regehr, 2009) generate a repertoire of responses in the output of the granule cells. We will refer to these outputs as the *cortical basis*. Such cortical basis, relayed by the parallel fibers, serve to modulate the output of Purkinje cells. Since parallel fibers can directly excite a Purkinje cell or inhibit it by disynaptic inhibition through the molecular layer interneurons, then a same parallel fiber can have either a positive or a negative effect in the projecting Purkinje cell. This allows the weights applied to the cortical basis in the model to have positive or negative values (Porrill & Dean, 2008). Such weights model the gain of synaptic contacts made by the parallel fibers with inhibitory interneurons and Purkinje cells. Their values are adjusted according to the teaching signal provided IO activity, that reaches the cerebellar cortex through the climbing fibers.

This implementation of the adaptive filter model includes the following novelties:

1. we use a random spike generator model for the IO with a rate of 1 Hz, consistent with the range of low firing rates observed in vivo Eccles et al. (1967);
2. we set a series of parallel microcircuits, each one with its own IO component;
3. we collapse the cerebellar cortex and cerebellar nuclei together.

The total output of the cerebellar controller is produced averaging the output of all the microcircuits.

Cortical basis. To generate the cortical basis we convolve the signal coming through the mossy fiber pathway with two exponentials. In such a way, the response of each basis to a unitary pulse resembles an alpha function. The time constants governing the exponentials are randomly drawn from two flat distributions (a fast time constant, τ_r, ranging from 2 to 50 ms and a slow one, τ_d, ranging from 50 to 750 ms). The first set of fast time constants control the rise of the basis and the second, the decay. These values are within the physiological range of the time constants of the slow currents in the granular layer, e.g., slow spillover inhibitory currents (Crowley et al., 2009; Hamann, Rossi, & Attwell, 2002; Rossi & Hamann, 1998).

Given an input $m(t)$, the output of a cortical basis $p_j(t)$ is generated according to the next difference equations:

$$p_j^r(t) = \gamma_j^r p_j^r(t-1) + m(t-1)$$

$$p_j^d(t) = \gamma_j^d p_j^d(t-1) + p_j^r(t-1)$$

$$p_j(t) = \sigma_j[p_j^d(t-1) - \theta_j]^+$$

where j indexes a particular basis. p_j^r and p_j^d compute a convolution and, informally, each one governs the rise and decay of the p_j basis, respectively. They are controlled by the persistence factors γ_j^r and γ_j^d, which are computed from τ_r and τ_d (see above). The third equation adds a non-linearity (a threshold θ_j) allowing to produce steeper responses (note that $[x]^+ = \max(x, 0)$). Such threshold was set to 0.7 times the maximum value attained by p_j^d when $m(t)$ carried a spike during a single time-step. Finally, all basis are scaled by σ_j such that their maximum amplitude is fixed to 1.

The output of each microcircuit is obtained by a linear combination of its current components:

$$C_i(t) = [\mathbf{p}_i(t)\mathbf{w}_i(t)^T]^+ \quad (1)$$

where i indexes a particular microcircuit, $\mathbf{w}_i(t)$ is the vector of weights and $\mathbf{p}_i(t)$ is the vector of basis. As in Lepora et al. (2010) we clipped the output of the adaptive filter to remove negative values.

The weights are updated using the de-correlation learning rule:

$$\Delta w_{ij}(t) = \beta\, e_i(t)\, p_{ij}(t-\delta) \quad (2)$$

where β controls the learning rate and $e_i(t)$ is the error signal of the microcircuit i, computed from the output of the IO (see below). δ corresponds to the latency of the error feedback (Miall, Christensen, Cain, & Stanley, 2007). This parameter establishes when the action, that caused or could have prevented the current error, occurred or should have occurred. For this, δ resolves the temporal credit assignment problem (Sutton & Barto, 1998). In practice, this parameter sets up the relative timing of the CR relative to the US. For this, it has to account not only for the delays in transmission of the signals, but also for the latency associated to the action execution. In our case, given the temporal dynamics of the temporal basis, turns last on the order of 100 ms, making such a value a sensible choice for δ.

2.2. Inferior olive

Since this work focuses on the NOI, we found it relevant to reproduce in our model, a distinctive trait of the IO physiology: its low spontaneous firing rate. To this end, each IO component generates randomly and independently spikes with a *Poisson* firing rate of 1 Hz. Having a simulation time step orders of magnitude below this rate, we generate the Poissonian firing pattern converting that rate into a baseline firing probability, b. The baseline probability is then modulated at each time step both by the behavioral error signal, $E(t)$, and by the cerebellar command of the i-th microcircuit, $C_i(t)$, yielding the probability of spike in the IO at each time step, $P_i^{\mathrm{IO}}(t)$, according to:

$$P_i^{\mathrm{IO}}(t) = b + E(t) - k_c C_i(t - \delta)$$

where k_c is the gain of the NOI. The last term of this equation implements the NOI, subtracting the cerebellar output to the probability of firing in the IO. Note, however, that δ, the same parameter that controlled the temporal eligibility trace in Eq. (2), is used here to delay the effect cerebellar output into the IO. Thus, as in nature, we set up a NOI with a long latency.

2.3. Error signal

To compute the internal error signal term in Eq. (2) each microcircuit compares the instant IO activity $r_i^{\mathrm{IO}}(t)$ and the baseline $\tilde{r}_i^{\mathrm{IO}}(t)$ as follows

$$e_i(t) = r_i^{\mathrm{IO}}(t) - \tilde{r}_i^{\mathrm{IO}}(t)$$

where $\tilde{r}_i^{\mathrm{IO}}(t)$ is computed by exponential averaging with a time constant of 10 s.

2.4. Scenario

We tested our model in a simulation where a robot performs a collision avoidance task. The simulated robot has a cylindrical shape and incorporates two sensors: a proximity sensor that detects the presence of a wall and a rudimentary visual sensor that allows the robot to detect a color mark in the ground. At every trial, as the robot traverses the track, it passes over a green mark in the ground (cue) and after a fixed timed interval, determined by the distance between the cue and the turn, detects the proximity to the wall and turns (Fig. 3). Note that with this scenario the robot avoidance task is analogous to an eyeblink AL experiment. We already exploited a similar analogy in Hofstotter, Mintz, and Verschure (2002).

To remark that this is a scenario of AL and not of classical conditioning, we use a different nomenclature. Instead of a US, an error signal E is relayed by the proximity sensors. E depends in the physical closeness, p, between the robot and the wall, that is computed relative to the radius of the robot r as follows

$$p = \left[1 - \frac{|\mathbf{c} - \mathbf{x}| - r}{r}\right]^+$$

where the vector $\mathbf{x}$ vector contains robot position, $\mathbf{c}$ the closest point in the wall and the operator $\|$ computes the vector norm. Note that the value of p ranges between 0 and 1. Then E is generated multiplying p by the approaching angle:

$$E = [\cos(\dot{\mathbf{x}}, \mathbf{c} - \mathbf{x})\ p]^+$$

where $\dot{\mathbf{x}}$ is the robot velocity. E can only have positive values, since negative ones imply that the robot moves away from the wall.

Instead of a CS we have a *cue* signal, which is a binary event fired when the robot crosses over the green line. The reactive response, similar to the unconditioned response is denoted as R, and the adaptive response, which is generated by the cerebellum and is similar to the CR, is denoted as C. Both responses induce a turn in the robot trajectory by controlling the angular velocity and are added to generate the total turn at each time step. R is generated by multiplying E by the gain of the reflexive controller (k_r).

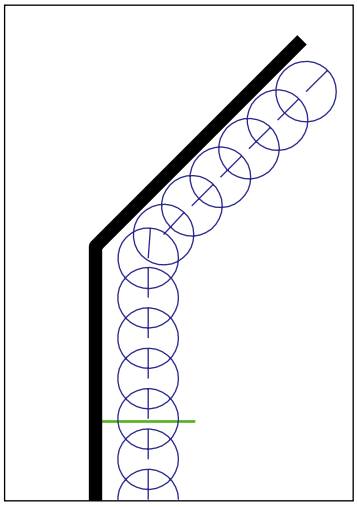

Fig. 3. A sample trial. Trajectory of the robot during a trial, replotted every 50 time steps. The green mark in the ground is the cue signal and the thick black line is the wall. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Results

3.1. Acquisition of a response

Before discussing the experimental results, we highlight a key difference between the obstacle avoidance task we use and AL of the eyeblink reflex. From the sensory prediction view, in the eyeblink paradigm the target function for the cerebellar controller to match – the sensory response to the US – varies in amplitude as the adaptive response evolves trial by trial; the noxiousness of the airpuff diminishes as it reaches a more closed eyelid. In the collision avoidance setup both the amplitude of the US – the proximity to the wall – and its timing vary. Any turn, insufficient as it might be to avoid hitting the wall, delays the collision, increasing the interval between cue and collision. Given this contingency, it is difficult to predict for a track configuration the optimal timing of the turn. However, we can predict that once the responses become stable, the relative timing of C and R will be determined by the delay of the NOI (δ, see Methods).

We display the results of a representative simulation where we use a track with a turn of 30° and the delay between cue and the onset of the proximity signal is of 300 ms. To maintain the analogy with conditioning we refer to this time interval as the Inter Stimulus Interval (ISI). After a few trials the controller produces adaptive responses, that during the remaining of the simulation are slowly adjusted both in timing and amplitude (Fig. 4(A)). The final responses are narrower and peak later than the initial ones (Fig. 4(C)). All adaptive trajectories are smoother than the purely reactive one, but within the adaptive trajectories we observe that the robot turns faster and closer to the wall in the last trials (Fig. 4(B)) and that by the end of training the adaptive response provides a higher proportion of the total response (Fig. 4(D)).

3.2. Adaptability of the responses in time and amplitude

We show that the controller adapts to different turn angles, R_0, (Fig. 5(A) and (B)) and ISIs (Fig. 5(C) and (D)). For a same ISI,

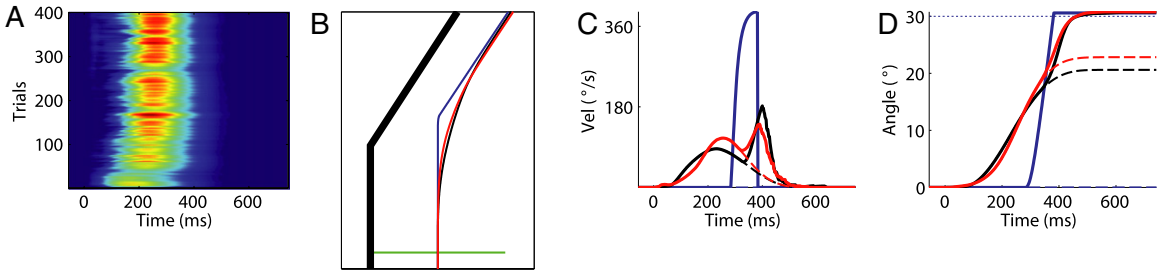

Fig. 4. Single simulation. A. Raster plot of the adaptive responses (C) trial by trial. B. Three sample trajectories plotted on the track: trial 1 (blue), mean of trials 51–100 (black) and trials 351–400 (red). The cue signal (green line) and the walls of the track are shown (thick black line). C. Profile of the total response, $C + R$, (solid lines) for the same data in B. The dashed lines separate the adaptive component, C, (below) from the reactive, R, (above). D. Cumulative responses. The dotted line marks the target response (R_0). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

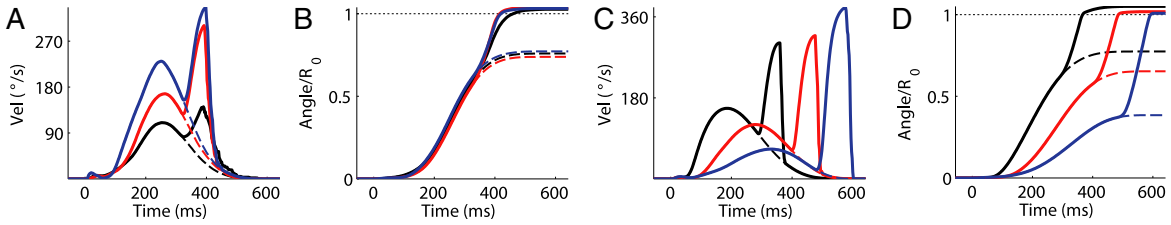

Fig. 5. Adaptability to different amplitudes and timings. A. Mean responses in the last 50 trials for three different target amplitudes: 30° (black), 45° (red) and 60° (blue). The dashed lines separate the adaptive component, C, (below) from the reactive, R, (above) B. Cumulative responses scaled to the target amplitude (R_0). Same data as in A. C. Adaptive and reactive responses for three different timings: 230 ms (black), 330 ms (red) and 430 ms (blue). Data displayed and trial selected as in A. D. Cumulative responses scaled to the target amplitude (R_0). Same data as in C. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

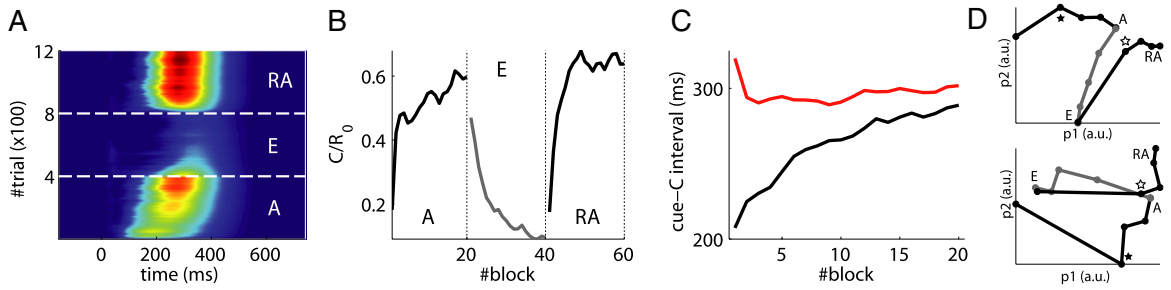

Fig. 6. Acquisition, extinction and reacquisition. A. Raster plot of the adaptive response C during acquisition (A), extinction (E) and reacquisition (RA). Warmer colors for higher amplitude. B. Evolution of the ratio of adaptive to total response (C/R_0). Blocks of 20 trials. C. Evolution of the timing of the adaptive response C during acquisition (black) and reacquisition (red). The timing of C is computed as $\frac{\sum C(t)t}{\sum C(t)}$. D. (Up) Trajectories of the weights during learning (First two principal components). The labels identify the end of each stage. The solid star marks the position after 100 trials of acquisition and the empty star, after 100 trials of reacquisition. (Down) Trajectory of the evolution of the adaptive responses (C). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the ratio of the adaptive response to R_0 (C/R_0) remains constant regardless of R_0 (Fig. 5(B)). This is not the case for different ISIs: the larger the distance from cue to collision, the minor the ratio of the adaptive response (C/R_0) (Fig. 5(D)). In short, as in nature, performance decreases as the ISI increases.

3.3. Control of the IO over the dynamics of the learning

We validate that the controller acquires and extinguishes actions according to changes in the stimuli contingencies. For this, we run a protocol with acquisition, extinction and reacquisition stages. For the acquisition and reacquisition we use the same track, with a turn of 45° and an ISI of 300 ms, and for extinction we use a straight track.

In the results we observe that, after the regular acquisition, the response is erased with extinction training (Fig. 6(A) and (B)). Remarkably, during reacquisition both the amplitude of the response and the timing are more rapidly adjusted than in acquisition (Fig. 6(C)). We analyze this result in detail since it goes beyond the expected properties of this adaptive filter implementation. In short, fast reacquisition might occur because after extinction the weight configuration does not return to the initial one, but to a closer configuration in the weight space (in this case a 25% closer). This relative proximity is not totally explained by the trace of the adaptive response that still remains after extinction, since the magnitude of this response, measured as area under response, is only a 17% of magnitude prior to extinction. In other words, extinction affects more the overt behavior than the underlying memory. If we only consider the first two principal components describing both the trajectory of the weights and the evolution of C during the experiment, the same results hold (Fig. 6(D)). However, by visual inspection we also appreciate that after just 100 trials of reacquisition, both the weights and C are very close to their configurations at the end of the previous acquisition, suggesting a faster (more direct) gradient descent during reacquisition.

3.4. Behavioral role of the NOI

We check how the different values of k_c affect the behavior. We use a track with a 45° for acquisition, and afterwards a straight

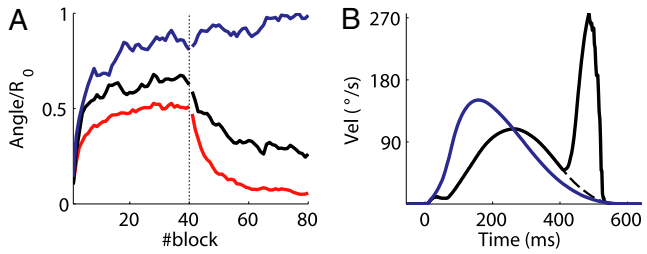

Fig. 7. Effect of the nucleo-olivary inhibition loop. A. Accumulated adaptive response per trial proportional to the target response (R_0) in the two stages of the experiment (acquisition and extinction) for three gains of the NOI (k_c); values k_b^{-1} (red), $k_b^{-1}/4$ (black) and 0 (blue). B. Mean final responses C (solid) and $C+R$ (dashed) for two of the models. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

track for extinction. We see that the higher the value of k_c, the smaller the adaptive response but the faster and more effective the extinction (Fig. 7(A)). As expected, after removing the NOI no extinction is observed. More interestingly, the NOI is required for the precise timing of the responses (Fig. 7(B)). Since the basis generated in the cerebellar cortex are not optimal to adjust the timing, part of the initial response has to be deleted or unlearned (Fig. 4(C)). This latter process relies on the NOI. Thus, in the absence of the NOI, the initial response, poorly timed, is maintained as long as it avoids the collision (Fig. 7(B)).

3.5. Transference of responses between layers controlled by the gain of the NOI

Finally, we test how the different values of k_c balance the adaptive and reactive responses. The values of k_c are chosen in proportion to the inverse of the gain of the reflexive controller k_r. First, we see that for a same ISI, the ratio between the adaptive and reactive actions is set by k_c independently of the angle of the turn (Fig. 8(A)). This is, at the end of acquisition simulations with a same k_c have a same proportion of the initial response transferred to the adaptive layer. Next, in a same track with a turn of 45° and an ISI of 300 ms, we run simulations varying the value of k_c between $2k_r^{-1}$ and $2^{-6}k_r^{-1}$. Higher values of k_c result in a minor transfer of the response to the adaptive layer (Fig. 8(A)). When k_c is equal to the inverse of k_r so that the feedforward connection of the reflexive controller and the feedback connection of the NOI apply inverse transformations, then the final response is roughly equally distributed between the adaptive and the reactive layers. As k_c becomes smaller, the maximum proportion of transfer achieved is of 0.8, i.e., the anticipatory action executes an 80% of the total turn. However, minimizing the reactive action by transferring more control to the adaptive layer comes at the expense of the timing accuracy of the anticipatory action. Only the smaller values of k_c yield a C closely timed to the ISI, with C anticipating R by 100 ms. This is, for small values of the k_c the temporal arrangement of C and R reflects accurately the latency of the NOI δ. Summing up, k_c trades off the amount of anticipatory control against the timing accuracy.

4. Discussion

It is suggested that one of the functions of the cerebellum is to replace reflexes by anticipatory avoidance actions (Wolpert, Miall, & Kawato, 1998). In this paper, we proposed that the NOI might allow such a replacement to be partial, resulting in a cooperation between the reactive and adaptive layers of control. To test this hypothesis we built a computational model including a reflex controller – the reactive layer – and a cerebellar controller – the adaptive layer – implementing the latter along the lines of the cerebellar adaptive filter theory (Dean et al., 2010; Fujita, 1982). This architecture was then applied to a simulation of a robot collision avoidance task. The results confirmed that the strength of the NOI inhibition determined the degree of replacement of the original reactive turn by an anticipatory adaptive one. Moreover, since the question of whether the cerebellum performs motor control or sensory prediction in classical conditioning still remains open (Lepora et al., 2010), such results advocate for the possibility of both roles coexisting in cerebellar function.

The originality of our approach lies in the inclusion of the NOI in a motor control scenario. On one hand, previous computational simulations of the cerebellum that included the NOI have only been applied to classical conditioning paradigms, where the CR affects the US signal reaching the cerebellum through the NOI, but not behaviorally (Medina & Mauk, 2000). This left open the question of how the NOI might interact with a behavioral avoidance response. On the other hand, simulations that used a cerebellar model in motor tasks, such as navigating a robot through a curved path (McKinstry, Edelman, & Krichmar, 2006), lacked the NOI feedback projection.

Multiple roles are attributed to the inhibitory projections from deep nuclear cells to the IO (Bengtsson & Hesslow, 2006): controlling the tonic activity of the whole olivo-cortico-nuclear circuit (Andersson & Hesslow, 1987; Demer, Echelman, & Robinson, 1985), regulating the teaching signal provided by the climbing fibers (Jirenhed et al., 2007; Medina et al., 2002) and also adjusting dynamically the coupling between of olivary cells through gap junctions (De Zeeuw et al., 1998; Llinas & Welsh, 1993). In this paper we have only dealt with the control of the teaching signal.

4.1. NOI in timing and optimality of the cortical basis

Comparing avoidance responses at initial and later stages of learning, we see that the peak of the responses is delayed with learning until a more optimal timing is achieved. This operation requires suppressing the early part of the initial responses. But to correct these initial poorly timed responses, that are already effective from a behavioral point of view, the cerebellum has to detect the mismatch between its positive output and the shape of the error signal.

However, if we had used a set of basis that was optimal for the response generation (e.g., a set of Gaussian functions of constant amplitude but increasing delays as in Lepora et al. (2010)) then it would have not been necessary to suppress early components of the response. Interestingly, in classical conditioning the time of CRs is adjusted in a similar manner during learning in mice: training delays the peak of the CR until it matches the US onset (Koekkoek et al., 2003). Thus, using the adaptive filter stance for speculation, we can infer that the basis produced in the cerebellar cortex might not be optimal for the timing of delayed eyeblink responses.

4.2. NOI and extinction

The necessity of the NOI for extinction has already been shown in studies of classical conditioning both computationally and in physiology (Medina & Mauk, 2000; Verschure & Mintz, 2001). However we highlight that in the case of the learning of avoidance behavior, a controller both able to acquire and erase adaptive responses can function in a completely autonomous manner.

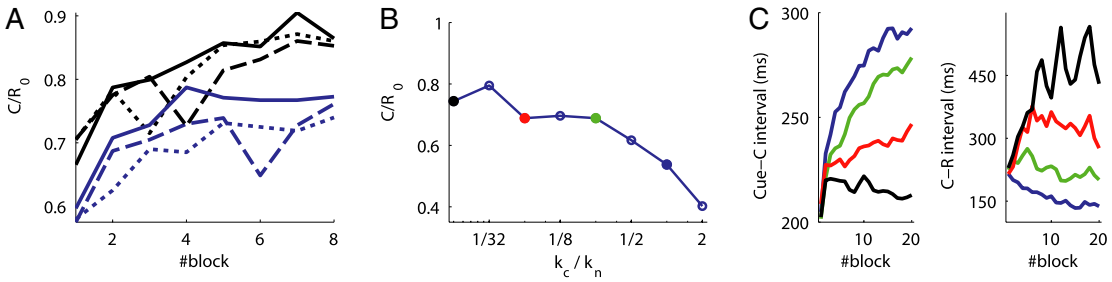

Fig. 8. Evolution of the proportion of the adaptive response (C/R_0) for different gains of the NOI (k_c) and total angle of the turn (R_0). A. Proportion of the adaptive over the total turn (C/R_0). Angles are 30° (dotted lines), 45° (dashed lines) and 60° (solid lines). We use gains of the NOI (k_c) of value $k_r^{-1}/2$ (blue) and $k_r^{-1}/4$ (black). B. Final values of C/R_0 as a function of k_c in a track with a turn of 45° and an ISI of 300 ms. C. Timing of C. Time between cue and C. Left. Time between C and R. The values of k_c correspond to the points with the same color in B. The timings are computed as in Fig. 6. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.3. Plausibility of a graded error signal

Our solution requires a graded error signal. We assume that the residual error either does not require correction or that the reactive response prevents any noxious consequences. In our simulations, even though we have not referred to the concepts of harm or noxiousness in the error signal, we have modeled the latter case; the residual error still induces a turn, a reactive response. It is our view that there are AL paradigms that provide a graded error signal. For instance, in eyeblink AL, the activation of the cornea receptors by the airpuff will be more harmful than the activation of the eyelid receptors. And obviously, even if the noxious US is completely avoided, i.e., even if the eyelids are totally closed protecting the cornea, this does not preclude the eyelid receptor to sense the airpuff.

4.4. Cerebellum as a forward model

We have shown that in the context of adaptive reflexes, interpreting the cerebellum as a forward model helps to define the goal of learning, as it to be the exact prediction of the sensory input. However, since in this architecture the same signal supports the sensory prediction and the motor control, this solution implies that the CR amplitude and the residual US signal are proportional. This proportionality might be non-intuitive. For instance, in eyeblink AL, we might expect the successful CR to be defined in terms of degree of eyelid closure independently of the intensity of the airpuff. In contrast, our model predicts that the CR amplitude at the asymptotic level depends on the US intensity. This prediction could be easily validated experimentally, or by a further analysis of existing data.

However a question stands regarding the suitability of the controller architecture here presented. Since there is evidence of cerebellar regions acting as forward models (Miall et al., 2007), why do we not adhere to the standard forward model-forward controller architecture (Miall et al., 1993)? Quite simply, AL is not a feedback control task and for this, a forward model-feedback controller architecture cannot be applied. For instance, in collision avoidance, we are not trying to act faster after the collision by predicting it, but to avoid the collision altogether. In our case, only the reflex controller performs feedback control, but the cerebellar controller uses the feedback information to improve performance in future trials.

4.5. Role of the inferior olive

Our computational model can acquire correct responses with the IO sustaining a physiologically correct rate of activity of 1 Hz. However, such a low firing rate slowed learning and caused performance of the individual microcircuits to fluctuate (data not shown). A problem that we practically by-passed by a adding independent IO cells to the model.

However, within each microcircuit, the low firing rate of the IO limited the gain of the NOI. In our implementation the error information resides the spiking probability of the IO (P^{IO}, see Methods). Therefore, whenever the NOI drives this probability below 0, information regarding the magnitude of the error is lost. For this, it is hard to see how can be beneficial to have a teaching signal with such a low firing rate. One view is that the output of the IO, besides acting as a teaching signal, might play other roles wherein low firing rate is beneficial. This role should be supported by the specific response of Purkinje cells to the climbing fiber signal, namely, the complex spikes. A second explanation is that the low rate of complex spikes might prevent them to interfere with the information carried at a much higher firing rate by the simple spikes.

4.6. Fast reacquisition

Although not directly related to the claim of this paper, we also reported that our model reproduces the fast reacquisition observed after extinction (Jirenhed et al., 2007). Previous computational studies reproduced fast acquisition by including different learning mechanisms with different learning rates (Medina, Garcia, & Mauk, 2001; Porrill & Dean, 2008). In our model fast reacquisition stems for a non-linearity in the cerebellar adaptive filter and not from the interaction of different learning mechanisms. Indeed, by the rectification applied to the cerebellar output (Eq. (1)) extinction can be achieved without returning to the original configuration of null weights, i.e., any configuration yielding a negative output before the rectification is also effective. Such a configuration might be closer in *weight space* to the weight configuration that triggers the CRs, for which at reacquisition returning to this configuration might be achieved faster.

4.7. Means to adjust the gain of the NOI

As we manually set the gain of the NOI (k_c) to different values, a next step would be to devise a heuristic or learning rule to set its gain adaptively. On one hand, our controller has to produce an adaptive avoidance response managing to decrease the error signal under a safety level. In addition, we know that, greater the gain of the NOI, the smaller the adaptive response produced. For this, an heuristic should decrease such gain to allow a sufficient transfer of control from the reactive to the adaptive layer. On the other hand, the smaller the NOI, the slower the extinction of no longer adaptive responses. Considering that the drive for rapid extinction might be related to the cost of unnecessary avoidance actions (for instance, closing the eyelids to protect the eyeballs results in a momentary loss of potentially relevant visual input), this cost function could

be used to learn an optimal k_c. In any case, a model solving this issue would have to hypothesize which extra-cerebellar structures evaluate this cost, and how does their information reach nucleo-olivary circuit.

4.8. Conclusion

We have applied a cerebellar controller including the NOI negative feedback-loop to the acquisition of an adaptive reflex in a collision avoidance task. Within that domain, we have shown (a) that we can simultaneously interpret the output of the cerebellar controller as a sensory prediction and as a motor control signal and (b) that the degree of transference of the initial reactive response into an adaptive response after learning is determined by the gain of the NOI. These results require the error signal, the US, to be graded. Moreover, under these conditions the NOI controls the temporal accuracy of the responses and allows the extinction of no longer adaptive responses. With this work, we have shown that from the cerebellar microcircuit studied in classical conditioning we can obtain a functional controller for AL. The next question now is whether once more realistic motor plants are taken into consideration, such as the muscles controlling the eyelids, this architecture is complete or additional components are necessary to simultaneously meet the requirements of sensory prediction and motor control.

Acknowledgements

Work supported by eSMC FP7-ICT- 270212.

References

Albus, J. (1971). A theory of cerebellar function. *Mathematical Biosciences*, *10*(1–2), 25–61.
Andersson, G., Garwicz, M., & Hesslow, G. (1988). Evidence for a gaba-mediated cerebellar inhibition of the inferior olive in the cat. *Experimental Brain Research*, *72*(3), 450–456.
Andersson, G., & Hesslow, G. (1987). Activity of purkinje cells and interpositus neurones during and after periods of high frequency climbing fibre activation in the cat. *Experimental Brain Research*, *67*(3), 533–542.
Bengtsson, F., & Hesslow, G. (2006). Cerebellar control of the inferior olive. *The Cerebellum*, *5*(1), 7–14.
Christian, K., & Thompson, R. (2003). Neural substrates of eyeblink conditioning: acquisition and retention. *Learning & Memory*, *10*(6), 427.
Clark, R., & Squire, L. (1998). Classical conditioning and brain systems: the role of awareness. *Science*, *280*(5360), 77–81.
Crowley, J., Fioravante, D., & Regehr, W. (2009). Dynamics of fast and slow inhibition from cerebellar Golgi cells allow flexible control of synaptic integration. *Neuron*, *63*(6), 843–853.
Dean, P., Porrill, J., Ekerot, C., & Jörntell, H. (2010). The cerebellar microcircuit as an adaptive filter: experimental and computational evidence. *Nature Reviews Neuroscience*, *11*(1), 30–43.
Demer, J., Echelman, D., & Robinson, D. (1985). Effects of electrical stimulation and reversible lesions of the olivocerebellar pathway on purkinje cell activity in the flocculus of the cat. *Brain Research*, *346*(1), 22–31.
De Zeeuw, C., Hoogenraad, C., Koekkoek, S., Ruigrok, T., Galjart, N., & Simpson, J. (1998). Microcircuitry and function of the inferior olive. *Trends in Neurosciences*, *21*(9), 391–400.
Eccles, J., Ito, M., & Szentágothai, J. (1967). *The cerebellum as a neuronal machine*. Berlin: Springer.
Fujita, M. (1982). Adaptive filter model of the cerebellum. *Biological Cybernetics*, *45*(3), 195–206.
Gormezano, I., Prokasy, W., & Thompson, R. (1987). *Classical conditioning*. Lawrence Erlbaum.
Hamann, M., Rossi, D., & Attwell, D. (2002). Tonic and spillover inhibition of granule cells control information flow through cerebellar cortex. *Neuron*, *33*(4), 625–633.
Hesslow, G. (1986). Inhibition of inferior olivary transmission by mesencephalic stimulation in the cat. *Neuroscience Letters*, *63*(1), 76–80.
Hesslow, G. (1994). Inhibition of classically conditioned eyeblink responses by stimulation of the cerebellar cortex in the decerebrate cat. *The Journal of Physiology*, *476*(2), 245.
Hofstotter, C., Mintz, M., & Verschure, P. (2002). The cerebellum in action: a simulation and robotics study. *European Journal of Neuroscience*, *16*(7), 1361–1376.
Ito, M., Sakurai, M., & Tongroach, P. (1982). Climbing fibre induced depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar Purkinje cells. *The Journal of Physiology*, *324*(1), 113.
Jirenhed, D., Bengtsson, F., & Hesslow, G. (2007). Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace. *Journal of Neuroscience*, *27*(10), 2493.
Koekkoek, S., Hulscher, H., Dortland, B., Hensbroek, R., Elgersma, Y., Ruigrok, T., & De Zeeuw, C. (2003). Cerebellar LTD and learning-dependent timing of conditioned eyelid responses. *Science*, *301*(5640), 1736.
Lepora, N., Porrill, J., Yeo, C., & Dean, P. (2010). Sensory prediction or motor control? application of marr-albus type models of cerebellar function to classical conditioning. *Frontiers in Computational Neuroscience*, http://dx.doi.org/10.3389/fncom.2010.00140.
Llinas, R., & Welsh, J. (1993). On the cerebellum and motor learning. *Current Opinion in Neurobiology*, *3*(6), 958–965.
Mackintosh, N. (1974). *The psychology of animal learning*. New York: Academic Press.
Mapelli, J., Gandolfi, D., & D'Angelo, E. (2010). Combinatorial responses controlled by synaptic inhibition in the cerebellum granular layer. *Journal of Neurophysiology*, *103*(1), 250–261.
Marr, D. (1969). A theory of cerebellar cortex. *The Journal of Physiology*, *202*(2), 437.
Mauk, M., Steinmetz, J., & Thompson, R. (1986). Classical conditioning using stimulation of the inferior olive as the unconditioned stimulus. *Proceedings of the National Academy of Sciences*, *83*(14), 5349.
McKinstry, J., Edelman, G., & Krichmar, J. (2006). A cerebellar model for predictive motor control tested in a brain-based device. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(9), 3387–3392.
Medina, J., Garcia, K., & Mauk, M. (2001). A mechanism for savings in the cerebellum. *The Journal of Neuroscience*, *21*(11), 4081–4089.
Medina, J., & Mauk, M. (2000). Computer simulation of cerebellar information processing. *Nature Neuroscience*, *3*, 1205–1211.
Medina, J., Nores, W., & Mauk, M. (2002). Inhibition of climbing fibres is a signal for the extinction of conditioned eyelid responses. *Nature*, *416*(6878), 330–333.
Miall, R., Christensen, L., Cain, O., & Stanley, J. (2007). Disruption of state estimation in the human lateral cerebellum. *PLoS Biology*, *5*(11), e316.
Miall, R., Weir, D., Wolpert, D., & Stein, J. (1993). Is the cerebellum a smith predictor? *Journal of Motor Behavior*, *3*(25), 203–216.
Pavlov, I., & Anrep, G. (1927). *Conditioned reflexes*. Dover Pubns.
Porrill, J., & Dean, P. (2008). Silent synapses, LTP, and the indirect parallel-fibre pathway: computational consequences of optimal cerebellar noise-processing. *PLoS Computational Biology*, *4*(5).
Rescorla, R., & Wagner, A. (1972). A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning II: current research and theory* (pp. 64–99).
Rossi, D., & Hamann, M. (1998). Spillover-mediated transmission at inhibitory synapses promoted by high affinity [alpha] 6 subunit GABAA receptors and glomerular geometry. *Neuron*, *20*(4), 783–795.
Steinmetz, J., Lavond, D., & Thompson, R. (1985). Classical conditioning of the rabbit eyelid response with mossy fiber stimulation as the conditioned stimulus. *Bulletin of the Psychonomic Society*, *23*(3), 245–248.
Sutton, R., & Barto, A. (1998). *Reinforcement learning: an introduction, volume 28*. Cambridge Univ. Press.
Verschure, P., & Mintz, M. (2001). A real-time model of the cerebellar circuitry underlying classical conditioning: a combined simulation and robotics study. *Neurocomputing*, *38*, 1019–1024.
Wolpert, D., Miall, R., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*(9), 338–347.
Yamazaki, T., & Tanaka, S. (2007). The cerebellum as a liquid state machine. *Neural Networks*, *20*(3), 290–297.
Yeo, C., & Hesslow, G. (1998). Cerebellum and conditioned reflexes. *Trends in Cognitive Sciences*, *2*(9), 322–330.