A cerebellar population coding model for sensorimotor learning

Tianhe Wang*, Richard Ivry

Department of Psychology and Helen Wills Neuroscience Institute, University of California, Berkeley, California

Corresponding authors (*):

Tianhe Wang (tianhewang@berkeley.edu)

Abstract

The cerebellum plays a critical role in sensorimotor learning, and in particular using error information to keep the sensorimotor system well-calibrated. Here we present a population-coding model of how the cerebellum compensates for motor errors. The model consists of a two-layer network, one corresponding to the cerebellar cortex and the other to the deep cerebellum nuclei, where the units within each layer are tuned to two features, the direction of the movement and the direction of the error. We evaluated our model through a series of behavioral experiments that test sensorimotor adaptation across a wide range of perturbation schedules. The model successfully accounts for interference from prior learning, the effects of error uncertainties, and learning in response to perturbations that vary across different time scales. Importantly, the model does not require any modulation of the parameters or context-dependent processes during adaptation. Our results provide a novel framework to understand how context and environmental uncertainty modulate cerebellar-dependent learning.

Introduction

Humans are incredibly flexible in how we adapt our motor behavior across variable environments. We readily compensate for the added weight of a heavy winter coat when reaching for an object or adjust the force required as we sip on our morning coffee. The cerebellum is recognized as playing a key role in this adaptation process[1,2], utilizing errors as teaching signals to improve subsequent, similar movements[3,4]. This form of learning operates implicitly, automatically recalibrating the sensorimotor system without the need for awareness or drawing on cognitive resources[5–7]. The current paper aims to understand how this process is modified by context and environmental uncertainty.

Previous research has suggested that cerebellum-dependent learning is cognitively impenetrable, responding to error in a rigid manner even when the correction fails to improve task performance[5,8–11]. Moreover, unlike many learning processes, adaptation is not sensitive to the statistical properties of the perturbations[12,13]. However, this view of a rigid, inflexible system has been challenged by recent evidence showing that implicit adaptation is modulated by experience[14]. For instance, when participants are exposed to a previously experienced perturbation, the rate of relearning is slower than had been originally observed[15]. Not only does this result suggest a degree of flexibility in adaptation, but this context effect is opposite what is typically observed in studies of relearning: Across a broad range of task domains, relearning is typically faster[16–18]. This phenomenon, known as savings, is thought to reflect the reactivation of a residual memory. The rigidity and atypical effect of experience point to the need for considering the unique properties of the cerebellum in understanding how the processes of adaptation are modulated.

The basic principles of cerebellar-dependent error-based learning have been captured by the classic Marr-Albus model[2,19]. Purkinje cells (PC), the primary integrative unit in cerebellar cortex receive two types of input (Fig 1a). One source originates in the pontine nuclei which project to the granule cells of the

cerebellum. The axons of these granule cells, the parallel fibers, provide the primary input to PCs. This pathway provides contextual information, and in the domain of movement, is hypothesized to convey an efference copy of the motor command. PCs operate as a forward model, utilizing this input to predict the sensory consequences of the motor command [20,21]. The second source originates in the inferior olive. Here, the axons, the climbing fibers, provide a teaching signal, indicating a mismatch between the predicted and expected sensory feedback, that is used to update the forward model. Physiologically, activation of the climbing fibers induces long-term depression (LTD) of parallel fiber-PC (PF-PC) synapses, reducing the efficacy of similar input on PC activity.

Here, we expand on the Marr-Albus theory by incorporating some recent developments in cerebellar physiology to describe how the cerebellum is modulated by experience and environmental variability. First, recent studies have revealed a fundamental property of PCs: these cells are not only tuned to movement direction but also to the direction of error relative to that movement (Fig 1b)[22–24]. Second, learning is not confined to the cerebellar cortex; it also takes place in the deep cerebellar nuclei (DCN)[25–28]. By linking these two layers by positing connections between units that share similar tuning profiles, we develop a cerebellar population coding (CPC) model that can capture how different contextual factors affect sensorimotor adaptation.

To validate our model, we conducted a series of behavioral experiments employing various perturbation schedules. Specifically, we examined the effect of past experience, error uncertainty, error size, and variation in temporal dynamics in evaluating our model. Where relevant, we consider two alternative models that have been proposed to elucidate how context and environmental uncertainty modulate sensorimotor learning. One model is centered on the idea that the motor system develops context-specific motor repertoires and determines which repertoires to express based on contextual cues[29–31]. The other

model proposes that the learning rate of the adaptation system is modulated based on error history[22]. Our population-coding model provides the best fit in all the behavioral tests, even without positing context-dependent learning or having the capability to modulate learning parameters. As such, these results point to a parsimonious model of how the cerebellum supports sensorimotor adaptation across a broad range of contexts.

Results

Cerebellar Population Coding (CPC) model

The Marr-Albus model outlines how the cerebellar cortex can be viewed as a general error-based learning system, with an emphasis on how the anatomy and physiology are ideal for learning arbitrary associations[2,32]. The model has inspired many empirical tests spanning a range of sensorimotor behaviors[26,33–35]. This body of work has generally focused on the acquisition of individual behaviors in a constant environment, for example, testing visuomotor adaptation in response to a fixed perturbation. Here we extend the model, focusing on how learning is modulated when the environment is variable. A foundational idea for our model is inspired by a recent work showing how PCs in the oculomotor cerebellar cortex are simultaneously tuned to two kinds of information[22,23]. The first is movement direction, similar to that observed in many motor regions of the cortex and subcortex. The second is the direction of a visual error that arises during that movement (Fig 1b, c). Tuning in terms of movement direction is reflected in the simple spike activity of the Purkinje cells and tuning in terms of movement error is reflected in the complex spike activity of these cells. Importantly, because the two tuning profiles are in opposite directions, error-related activation will result in a change in the output to reduce that error.[22–24]

84 In the present study, we examine the implications of these tuning properties on cerebellar dependent
85 learning. To incorporate PC tuning into a learning model, we formulize the teaching signal, the complex
86 spike (CS) activity of a PC with a preferred direction of i ($0 \leq i < \pi$) in response to a movement error e (Fig
87 1d) as:

88 $$[1]\ CS_i^n = VM(\theta^e, i, s)\mathrm{F}(\rho^e)$$

89 where $VM(\ i, s)$ is the probability density function of a simplified circular (von Mises) distribution with a
90 mean of i and standard deviation of s. θ^e and ρ^e refer to the direction and the size of e, respectively, and
91 n is the trial number. F is a non-linear function to capture the well-established fact that learning rate does
92 not scale with error size[36,37]. Since this non-linear relationship is not a question focused by the current
93 study and we use a fix error size (Except Exp 6 & 8), $\mathrm{F}(\rho^e)$ was set as a 1. Following the Marr-Albus model,
94 the occurrence of a CS suppresses the strength of the parallel fiber input synapse (w) through long-term
95 depression (LTD):

96 $$[2]\ w_i^{n+1} = -lCS_i^n + f(w_\mathrm{o} - w_i^n) + w_i^n$$

97 where l ($l > 0$) and f ($0 < f < 1$) are the learning and forgetting rates, respectively, and w_o is the
98 baseline synaptic strength. Since the level of single spike (SS) activity will be greatest for cells coding a
99 movement direction opposite to the error, the modulation of synaptic strength will drive the next
100 movement in a direction that corrects for the observed error.
101
102 The preceding paragraph describes how parallel fiber synapses onto PCs are modified. A second
103 prominent site of plasticity is at deep cerebellar nuclei[25,38]. Lesion studies of eyeblink conditioning provide
104 one line of evidence indicating that some aspect of consolidated learning is centered in the DCN. Ablation
105 of the cerebellar cortex can completely block *de novo* cerebellar-dependent learning[26,39]. However, once
106 the learned behavior is established, it can persist after lesions to the cerebellar cortex even though the
107 kinematics are likely to be disrupted[40,41].

It has been hypothesized that this dissociation arises from a dual-effect of pontine projections to the cerebellum[42]. As described above, one pathway is via the polysynaptic projection to the cerebellar cortex (mossy fiber to parallel fiber to PC). The other is a direct, excitatory projection of the mossy fibers to the DCN. Importantly, PC and DCN neurons are organized such that they share the same tuning direction for movement[43]. We posit that learning at the DCN is gated by learning at the cerebellar cortex. Specifically, LTD at parallel fiber-PC (PF-PC) synapses will reduce inhibitory PC input to the DCN, resulting in long term potentiation (LTP) at the mossy fiber-DCN synapses (m) (Fig 1e):

$$[3]\ m_i^{n+1} = (w_{\mathrm{o}} - w_i^n)ß(m_{\max} - m_i^n) + \alpha(m_{\mathrm{o}} - m_i^n) + m_i^n$$

where ß and α are the learning rate and the forgetting rate of the DCN input synapse, respectively. The parameters m_{o} and $m_{\max}$ represent baseline and maximal synaptic strength, respectively. The latter constraint is based on empirical results showing that implicit adaptation saturates independent of the error size.

Considering the two sites of plasticity, DCN activity on a repeated trial following a movement error can be formalized as:

$$[4]\ DCN_i^{n+1} \propto m_i^{n+1} - \gamma w_i^{n+1}$$

where γ is a scale factor. The output of the population of DCN neurons will correspond to the change in movement direction in response to an error, a signal that can be used to adjust the movement. This can be expressed as: (Fig 1f):

$$[5]\ \boldsymbol{h}^{n+1} = -\varepsilon \sum_i \boldsymbol{v}_i DCN_i^{n+1}$$

where $\boldsymbol{h}^n$ is a vector representing the hand angle on trial n, $\boldsymbol{v}_i$ is a vector representing the tuning direction of unit i, and ε is a scale factor to transfer the neural activity into hand angle.

Putting together Eqs 1-5, we see that an error e will decrease parallel fiber input to PC (w) and increase mossy fiber input to DCN (m) (Fig 1d & e). Correspondingly, the net output of the cerebellum will provide a signal of the required change in movement direction to correct for the error (Fig 1f).

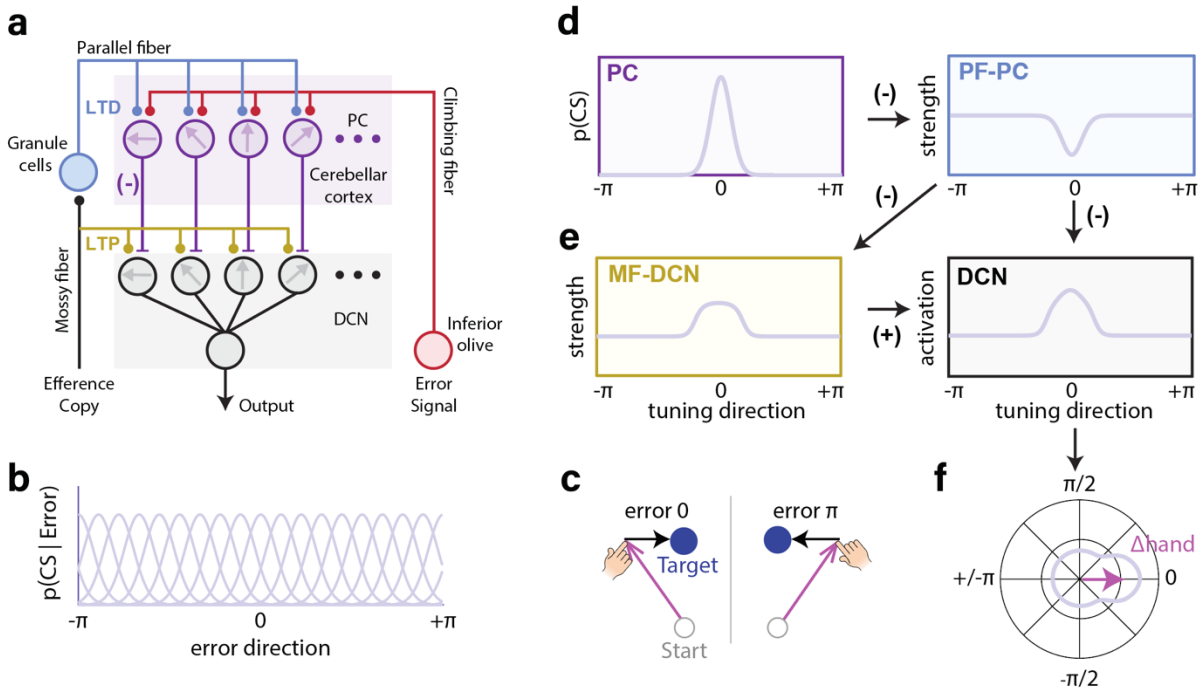


Fig. 1 Illustration of the CPC model. a) Structure of the cerebellar circuit incorporated in the CPC model. **b)** Each Gaussian-shaped curve represents the tuning function of a single Purkinje cell (PC) based on that cell's preferred error direction. For the simulations, we used 1000 units with preferred directions that covered 0-π in a uniform manner. **c)** Illustration of visual errors, with the direction of the error specified in polar coordinates. **d-e)** Model-generated adaptation in the cerebellar cortex (d) and deep cerebellar nuclei (DCN) (e). After experiencing an error in 0 direction, PC's with a preferred direction close to 0 will have high probability of generating a complex spike (CS) (d, left) which will result in long-term depression (LTD) for active synapses from granule cell inputs to that PC (d, right). During the preparation of the next movement, the strength of the input from the parallel fibers (PF) will decrease due LTD, attenuating the SS activity of the PC. Attenuation of PC output will result in long-term potentiation (LTP) at the mossy fiber (MF) input synapse to DCN (e, left). DCN activation is determined by the excitatory input from the MF and the inhibitory signal from the PC (e, right). The color of the frames corresponds to the color of structure in the Panel a. **f)** DCN activation

150 plotted in a polar coordinate. Activation across the population of cells results in a vector (purple arrow)
151 indicating the change in hand angle (Δ hand). Note that the vector points in the same direction as the error (c,
152 left), and thus serves to compensate for the error.
153

154 Clamp rotation task

155 In the following sections, we examine the CPC model, evaluating how well it captures a range of
156 phenomena observed during sensorimotor adaptation in various environments and under various training
157 schedules. In the empirical studies, we use a visuomotor rotation task in which the visual feedback during
158 a reaching task is limited to a cursor. To isolate cerebellar-dependent sensorimotor adaptation, we used
159 task-irrelevant clamped feedback in which the radial position of the cursor is locked to the hand, but the
160 angular position is fixed, shifted by a constant angle relative to the target (clamped)[5,42–44]. As such, the
161 angular position of the cursor is independent of the position of the participant's hand (Fig. 2a-b) and this
162 "error" remains constant across the perturbation phase of the experiment. Participants are fully aware of
163 the manipulation and instructed to ignore the feedback.
164

165 In Exp 1, we used a clamp with a fixed angle of 30° during a 100-trial training phase followed by a 60-trial
166 no-feedback "washout" phase (Fig 2c). As with prior studies using clamped feedback[5,44–46], we observed
167 prominent adaptation. When the perturbation was present, the reach angle shifted in the opposite
168 direction of the clamp and became relatively asymptotic by around the 30th reach. When the perturbation
169 was removed and feedback eliminated, the hand angle slowly shifted back towards the baseline direction.
170 These behavioral changes occur outside awareness[47]. Clamp-induced adaptation has all of the hallmarks
171 of implicit adaptation and, as with other forms of this type of learning, is dependent on the integrity of
172 the cerebellum[5,47].
173

The data from Exp 1 were used to determine the learning rate and forgetting rate for plasticity at the DCN, along with the scaling factor. The other two parameters of the model, the learning and forgetting rates for plasticity at the PCs, were empirically estimated in Exp 6 (see below). These parameters were fixed in the model simulations for the other experiments.

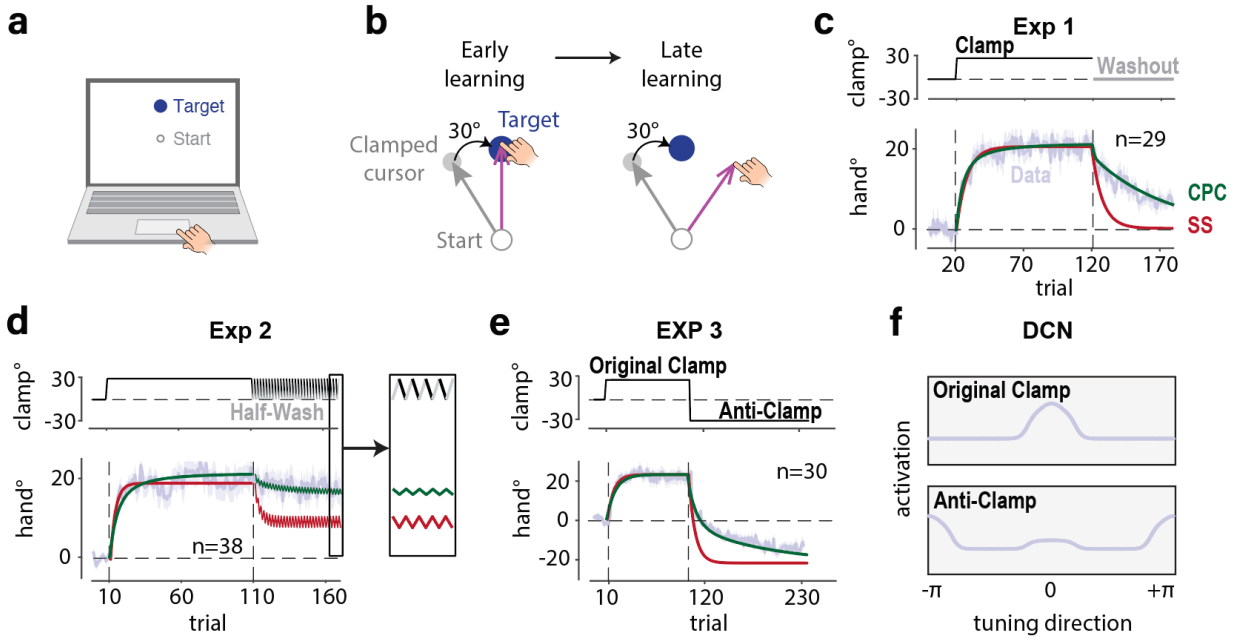


Fig. 2 Cerebellar population coding captures learning, forgetting, and anterograde interference during implicit adaptation. a) For online testing, stimuli are presented on the participant's laptop computer and movements are made on the trackpad. **B)** For clamped feedback, the angular position of the cursor is rotated by 30° with respect to the target, regardless of the heading direction of the hand. For a 30° clockwise clamp, the error direction remains at 0 on all trials. Similarly, for a -30° rotation (counterclockwise), the error direction would be invariant at π. **C)** Perturbation schedule (top) and results (bottom) for Exp 1. Time course of hand angle is shown in light violet. The CPC model provides a good fit in both the training and no-feedback washout phases. The state-space (SS) model using parameters fit from the training phase overestimates the forgetting rate in the washout phase. **d)** In Exp 2, half washout phase entails a 50/50 mix of clamp and no-feedback trials. Consistent with the CPC model, hand angle showed a small reduction whereas the state-space model predicts the hand angle will be reduced by 50%. **e)** We reversed the clamp direction during the training section of Exp 3. The behavioral results match the prediction of the CPC. **f)** Memory of the original perturbation (top row)

persists in the anti-clamp training phase, indicated by the activation of neurons tuned to 0 in the bottom row. This residual memory causes anterograde interference. Note that model all of the parameters were fixed in generating the predictions for Exps 2-3. Shaded area in c, d, e indicates standard error.

Failure of state-space models to provide parsimonious account of learning and forgetting during adaptation.

The CPC model provides an excellent fit to the learning function, including the washout period (Fig 2c). We recognize that this is not surprising given the number of parameters and relatively simple manipulation. However, this result stands in contrast to that obtained when these data are fitted with the most widely used sensorimotor adaptation model, the state-space model[44,48]. A key feature of the state-space model is that adaptation reaches an asymptote when the trial-by-trial effects of learning and forgetting cancel each other out (see Methods). While this will produce asymptotic learning, the state space model will predict a washout function that is much faster than empirically observed. The CPC model captures performance during both the acquisition and washout phases because the model includes a parameter specifying the upper boundary of adaptation, $m_{\max}$, assumed to reflect a limitation in neuroplasticity in the DCN.

To further compare the two models, we conducted a second experiment in which the post-training phase alternated between no-feedback and clamp trials (Exp 2, referred to as half-washout, Fig 2d). The state-space model predicts that the asymptote will drop to 50% because learning will only occur on 50% of the trials (feedback trials), a prediction that holds even in state-space models that posit learning at multiple time scales[18,49]. However, the asymptote showed only a slight decrease when clamped feedback was presented on 50% of the trials, consistent with the predictions of the CPC model (Fig 2d). These results highlight a major limitation in using a state-space model to capture implicit adaptation even when there is no manipulation of the learning context.

218

219 Anterograde interference

220 Having shown that the CPC model can capture the basic features of sensorimotor adaptation, we now
221 turn to phenomena in which adaptation is influenced by the experimental context. For units aligned with
222 a specific error, synaptic strength undergoes rapid changes due to the potent impact of complex spikes.
223 However, the recovery or resetting of these synapses during washout follows a relatively slow decay. As
224 such, at the population level, the net output of the system will be influenced by the persistent state of
225 units that were tuned to a recent perturbation. For example, due to this persistent state, the rate of
226 adaptation should be attenuated when the system is presented with a perturbation in the opposite
227 direction of a recently experienced perturbation. This effect is known as anterograde interference and has
228 previously been shown to occur during implicit adaptation[50–54].

229

230 To examine whether our model can quantitively predict anterograde interference using the parameters
231 measured from previous experiments, we used a task in which the sign of the clamp was immediately
232 reversed after an initial training block (e.g., 30° followed by -30°, Exp 3, see Fig 2e). The results showed
233 that the rate of adaptation was slower in response to the reversed clamp compared to the original clamp.
234 Indeed, the degree of attenuation closely matched the CPC model's prediction (Fig 2f).

235

236 Attenuation in relearning and no spontaneous recovery

237 Anterograde interference has typically been explained by positing context-dependent learning
238 mechanisms[30,55,56]. For example, the contextual inference (COIN) model assumes that the motor system
239 forms separate memories for different contexts and chooses which memory to use based on the inferred
240 context[29]. To account for the results of Exp 3, COIN would first build a memory for the 30° perturbation
241 and then a second, distinct memory for the -30° perturbation. Anterograde interference would arise

because the introduction of the -30° perturbation would lead to some degree of recall of the response to the initial perturbation. Over time, this would shift to a bias to recall the response to the second memory.

As shown in the previous section, the CPC model does not posit distinct memories for different perturbations; rather anterograde interference emerges from the dynamics of the CPC model. While CPC and COIN make similar predictions about anterograde interference, the two models make differential predictions on another memory recall phenomenon, spontaneous recovery. Spontaneous recovery refers to reappearance of a previously extinguished response, even when error information has not been re-introduced (e.g., no feedback phase). A paradigmatic design to elicit spontaneous recovery in sensorimotor learning studies would be to train participants with a perturbation in one direction, extinguish the adapted behavior by shifting the perturbation in the opposite direction, and then testing movements without feedback (Fig 3a). Spontaneous recovery refers to the fact that the initial movements during the no-feedback phase are in the opposite direction of the initial perturbation (Fig 3c left). By the COIN model, spontaneous recovery occurs because of recall of the original context during the no-feedback phase. In contrast, the CPC model predicts that spontaneous recovery will not occur when learning is restricted to the implicit system since the model does not have a mechanism for context-dependent memory (Fig 3c right).

In Exp 4, participants were trained with a 30° clamp in one direction for 100 trials and then presented with the opposite clamp for 15 trials (Fig 3a). Pilot testing had shown that this was sufficient to extinguish the shift in hand angle observed to the initial perturbation. The critical test was the subsequent 30-trials no-feedback block. At odds with the prediction of COIN, we failed to observe spontaneous recovery (Fig 3c, see Fig S1b).

A second point of contrast between the CPC model and context-dependent models such as COIN can be obtained by re-introducing the initial perturbation after the no-feedback washout (Fig 3a, relearning). The COIN model predicts that relearning should be faster (i.e., exhibit savings) because the system has stored a memory of the initial perturbation. The CPC model predicts that exposure to the opposite error during the washout phase will induce anterograde interference; as such, relearning will now be attenuated. Again, the results support the CPC model: Adaptation during the re-exposure block was slower compared to initial learning (Fig 3e; Fig S1a).

These results suggest that, for implicit adaptation, context effects such as anterograde interference are an emergent property of the system's inherent dynamics. A core feature of the CPC model is the absence of context-dependent memory, which we assume, is important for other forms of sensorimotor learning, such as those associated with action selection[31].

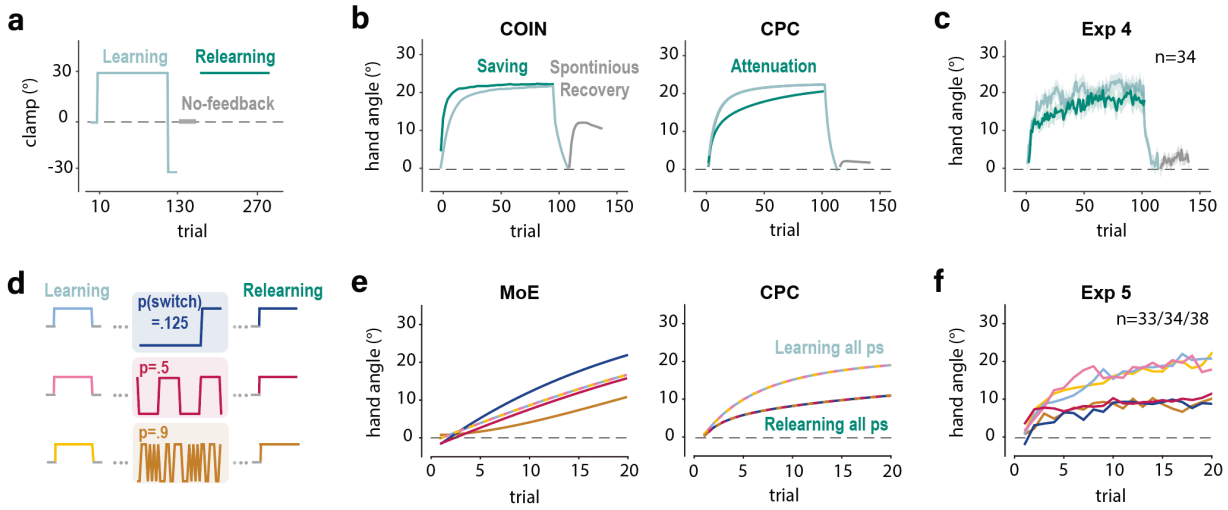


Fig. 3 Context as an emergent property of the CPC model. a) Exp 4 perturbation schedule. The learning and relearning phases are separated by an anti-clamp washout and no feedback phase to examine spontaneous recovery. The original perturbation is reintroduced in the second learning phase to test for savings. **b)** The COIN model predicts spontaneous recovery and savings; the CPC model predicts no spontaneous recovery and attenuation upon relearning. Note that, for visualization, the data from the relearning phase are plotted on top

of the original learning phase. **c)** Empirical results match both predictions of the CPC model. Shaded area in c indicates standard error. **d)** In Exp 5, the learning and relearning phases are separated by a variable phase in which the probability of a perturbation switch is manipulated between participants. **e)** The MoE model predicts that during relearning, the learning rate will be modulated by the prior switching rate (i.e., perturbation variability) whereas the CPC model predicts that the learning rate will not be modulated by switching rate. **f)** Empirical results are in accord with CPC model, showing attenuation and insensitivity to switching rate.

<u>Cerebellum-dependent learning is not sensitive to the consistency of errors.</u>

Context-dependent models confer a degree of flexibility on a learning process; in the case of COIN, the system is capable of storing multiple context-specific memories. An alternative form of flexibility is to allow the parameters of the model to change in response to context. For example, the Memory of Error (MoE) model assumes that the sensitivity of the system is modulated by error history[57]. Specifically, learning rate increases when the experienced errors are consistent and decreases when the experienced errors are inconsistent. Such a property is functionally useful in that the system will learn faster when the environment is relatively stable.

The CPC model does not provide such flexibility. Rather, we posit that the model parameters are fixed and experience-dependent changes in the response to an error arise because previous errors have transiently altered the state of the system. To compare the MoE and CPC models, we tested the response to a clamp with a fixed sign (e.g., 30°) before and after a block in which the sign of the clamp varied, with the switching probability set to 12.5%, 50%, or 90% (Fig 3d). The MoE predicts that the rate of relearning will be modulated by the switching frequency (Fig 3e left). However, consistent with the predictions of the CPC model, the rate of relearning was independent of the switching frequency (Fig 3e-f). Interestingly, relearning was markedly slower than the original learning (Fig S1c). This attenuation is another manifestation of anterograde interference resulting from the opposite errors experienced in the variable-clamp block.

These phenomenon have been previously explained as reflecting the operation of multiple processes that operate at different learning rates.[42,58] However, by the CPC model different learning rates can be an epiphenomenon of population coding rather than reflect the joint operation of multiple learning mechanisms (Fig S2). Cells with a preferred direction centered on the error direction will display relatively fast learning and quickly saturate. In contrast, cells with a preferred direction slightly misaligned with the error direction will learn slower due to the weaker climbing fiber input and take longer to saturate. The behavior change in movement direction is dictated by all of the units.

The CPC model provides a novel account of another phenomenon described in the adaptation literature, namely that behavior reflects the summed activity of multiple learning processes that operate at different rates. [42,58] An analysis of the CPC model shows that, in some cases, this can be an epiphenomenon of population coding rather than reflect the joint operation of multiple learning mechanism (Fig S2). Cells with a preferred direction aligned with the error will display faster learning and quickly saturate. In contrast, cells with a preferred direction slightly misaligned with the error direction will learn slower due to the weaker climbing fiber input and take longer to saturate. The net change in behavior is a composition of all activated units in the population.

<u>Labile and stable processes in cerebellum-dependent learning</u>

The preceding sections have focused on one critical feature of the CPC model, the population effects that emerge when individual units code both movement and error direction. We now turn to the other key feature of the model, that cerebellar-dependent learning involves an interaction between plasticity effects occurring in the cerebellar cortex and deep cerebellar nuclei, with the former gating the learning rate of the latter.

335

336 Our starting point here comes from an adaptation protocol in which the sign and size of the perturbation
337 randomly varies across trials. Consistent with previous studies[5,37,59], the response to clamped feedback in
338 this design (Exp 6) scaled for perturbations up to around 30°, and then saturated or even slightly
339 decreased for larger perturbations (Fig 4a). Importantly, when we empirically estimate forgetting (see
340 Methods), the retention rate is close to 0.5, indicating that about half of the learning from the previous
341 trial has been forgotten. This value stands in marked contrast with the empirically estimated retention
342 rate from designs in which the perturbation is fixed (Exp 1, 0.98). This difference is also found in previous
343 studies using either variable perturbations in a trial-by-trial design[36,37,60] or a fixed perturbation in a block
344 design[45] (Fig 4b,c).

345

346 The discrepancy in forgetting rates for blocked vs variable designs could be taken to indicate that
347 perturbation variability influences adaptation. However, given the results of Exp 5 (see also [61]), we
348 consider an alternative hypothesis, namely that implicit adaptation entails at least two processes that
349 operate on different timescales[18,58]. One process is labile, driving rapid changes that are weakly retained.
350 The other process is stable, producing changes at a relatively slow rate but with good retention (Fig 4h,
351 Fig S3).

352

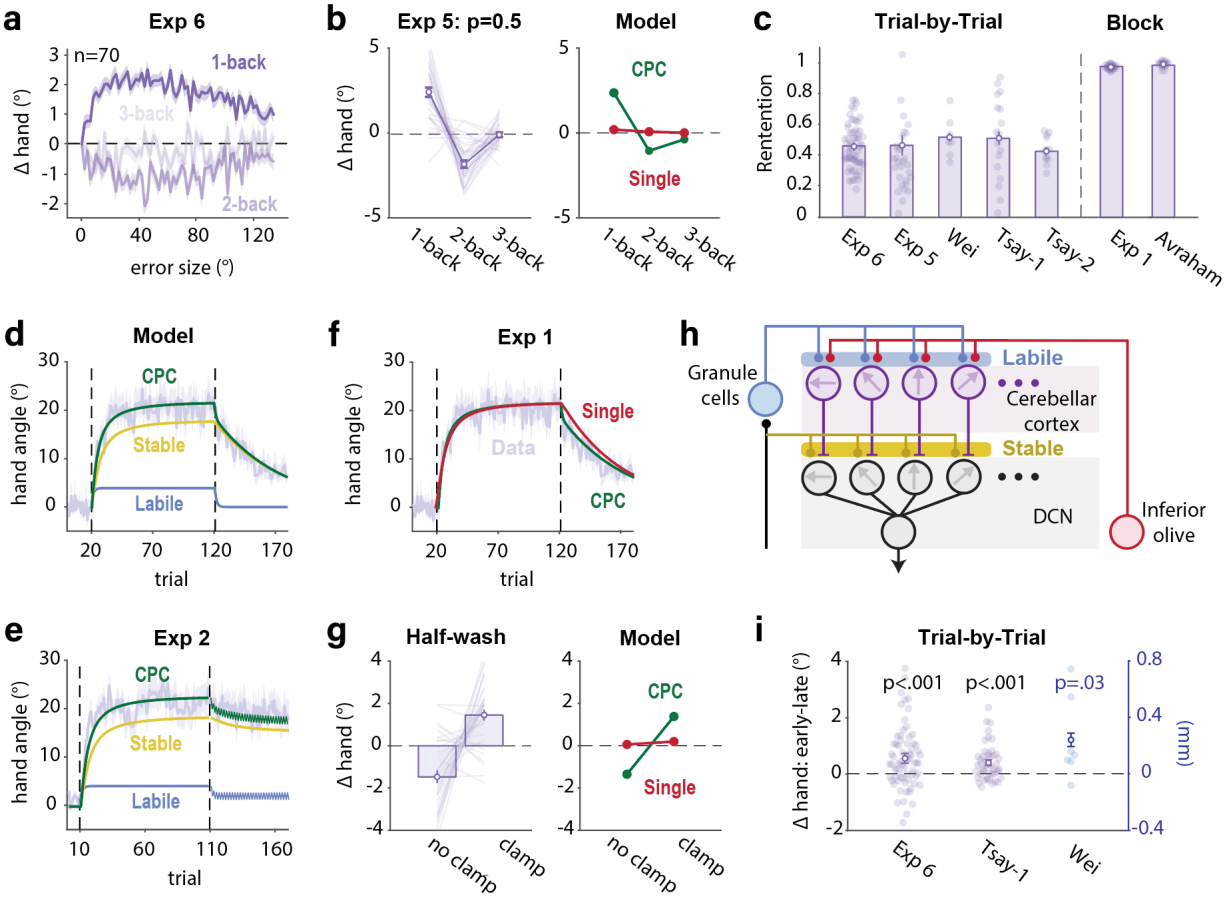


Fig. 4 Operation of labile and stable processes in cerebellum-dependent adaptation. a) Trial-by-trial change of hand angle (Δhand) as a function of the perturbation size on trial n-1 (1-back), n-2 (2-back), and n-3 (3-back) of Exp 6. **b)** Left: Δhand during variable phase of Exp 5 for the 50% switching condition. Right: The two-layer CPC model can account for the large change in hand angle observed in trial-by-trial designs whereas a single-layer model predicts a negligible change when the perturbation direction is varied. The forgetting rate can be empirically measured as the ratio between the Δhand of the 1-back and 2-back trials. **c)** Retention rate in the variable perturbation conditions is close to 0.50 whereas the retention rate in response to a fixed perturbation is close to 0.99. Re-analysis of data from Wei: [37]; Tsay-1: Exp 2 [60]; Tsay-2: Exp 3 [62], Avraham: Exp1 [45]. In all situations (trial-by-trial and blocked), we only included experiments tha used a single target. **d-e)** Simulated time course of the stable and labile processes in Exp 1 (e) and Exp 2 (f), along with their summed effect on behavior (CPC), and behavioral results. **f)** A single-layer model fails to account for the rapid forgetting observed at the start of the washout phase in Exp 1. **g)** The large Δhand in the half washout phase of Exp 2 (left) can only be accounted for by the two-layer CPC model (right). **h)** The labile process is hypothesized to produce LTD at the parallel fiber-PC synapse; the stable process is hypothesized to produce LTP at the mossy fiber-DCN synapse.

i) As predicted by the two-process CPC model, when exposed to a variable perturbation, the Δhand is larger in early training compared to late training. Shaded areas and error bars indicate standard error.

The dual operations of stable and labile learning processes is evident in other features of our data. In Exp 1, the retention rate changed over the course of the no-feedback block (Fig 4d), with large forgetting over the initial washout trials and much slower forgetting over the remainder of the block (also see Exp7, Fig S4). Fitting the results requires the joint operation of two forgetting rates operating at different temporal scales (Fig 4e). Moreover, in the half-washout condition of Exp 2, there is a considerable drop in hand angle after each no-feedback trial and a considerable increase after each clamp trial with a net result that the asymptote remains largely unchanged (Fig 4f & g). The large trial-by-trial change of hand angle near asymptote clearly suggests the operation of a labile process on top of a saturated stable process. Furthermore, with a variable perturbation, the learning rate is faster at the start of learning and decreases over the course of training (Fig 4i). Consistent with the CPC model, early learning reflects the operation of both the labile and stable processes, whereas during late learning, the stable process is saturated and performance changes are driven solely by the labile process (Fig S5).

<u>Learning of the stable process is gated by the labile process.</u>

Having seen that a two-process model is required to capture adaptation, we next consider the relationship between the labile and stable processes. In particular, do they make independent contributions to adaptation, or do they interact? As outlined in the Introduction, the CPC model proposes a specific form of interaction, namely that the stable process is gated by the labile process, with the former characterizing plasticity in the DCN and the latter plasticity in PCs (Fig 4h). This proposition is motivated by anatomical and physiological considerations: Anatomically, the output of the PCs provides the primary input to the DCN. Physiologically, learning in the cerebellar cortex occurs over a shorter time scale compared to the

394 DCN[25,38,42,63,64]. Given this two-layered network, we should find that the learning rate in the DCN is scaled
395 by the change in simple spike activity of the PCs.
396
397 As a first test of this hierarchical organization, we asked whether the stable and labile processes are
398 modulated in a similar way by perturbation size. Given the hierarchical assumption of the CPC model, we
399 should be able to measure the learning rate of the labile process with a variable perturbation design (Fig
400 5a) and use this to predict learning functions in response to fixed perturbations of different size.
401 Alternatively, if the stable and labile processes are independent, then we should not observe a similar
402 yoking. We considered two variants of an independent model (Fig 5b), one in which the learning rate of
403 the stable process is invariant across error size (FIX model) and one in which it is proportional to error size
404 (LINEAR model).
405
406 Consistent with the CPC model, using the estimates of the labile learning rate from Exp 6 (Fig 5a), we were
407 able to predict the learning functions in response to an invariant clamp of either 3° or 30° (Exp 8, Fig 5c).
408 The two alternative models (FIX and LINEAR) fail to account for the data (Fig 5d). Specifically, the learning
409 curves diverged during the early phase of learning before converging in late learning. When expressed as
410 the ratio of hand angle in response to the large perturbation relative to the small perturbation, we observe
411 a function that peaks early before dropping in a gradual manner (Fig 5e). We performed a similar analysis
412 on other data sets that involved a comparison of different error sizes[59] and found a similar pattern. In
413 contrast, we obtain a good fit with a model in which the labile process gates the stable process (Fig S6b-
414 c).
415

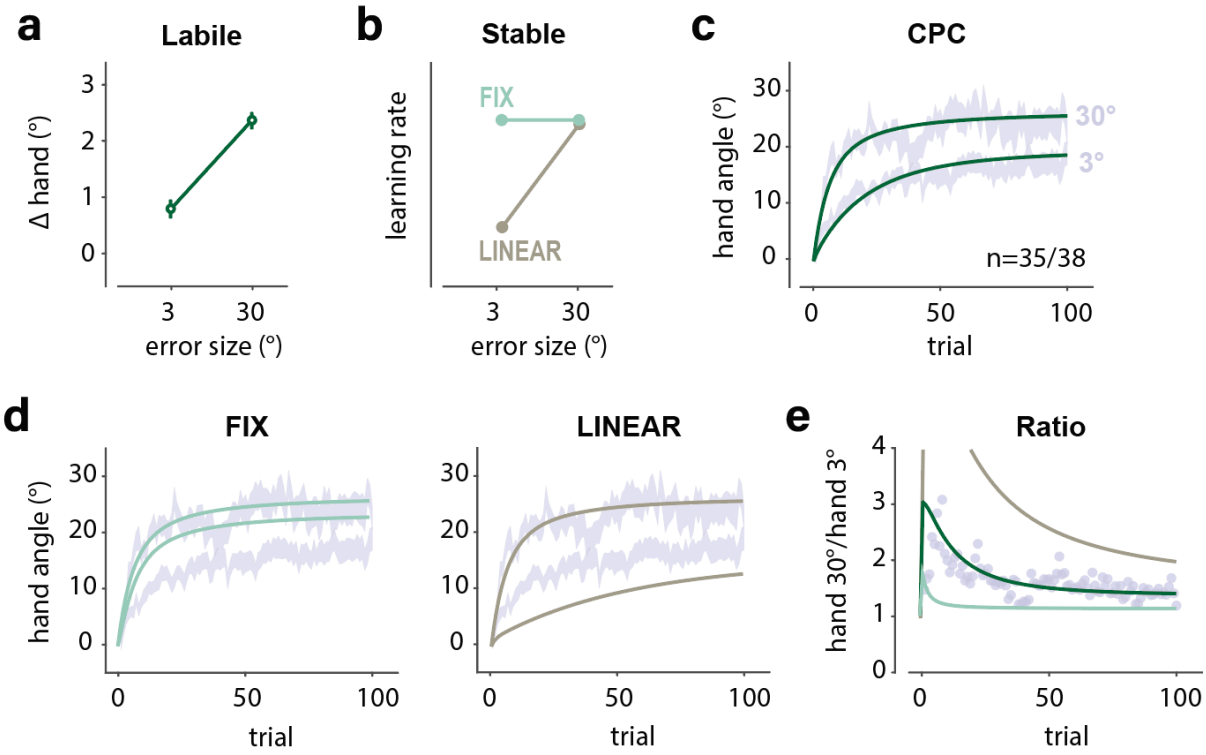


Fig. 5 Labile and stable learning rates are modulated in a similar manner by error size. a) The effect of error size on the labile process measured in Exp 6. For the CPC model, the ratio of the Δhand in a was used to estimate the learning rate of the stable process for the 3° condition (all other parameters are the same as in the 30° condition). **b)** Two alternative models in which the labile and stable processes are independent: In one, the learning rate of the stable process is invariant (FIX model) and in the other, the learning rate of the stable process is proportional to error size (LINEAR model). **c)** The CPC model (green) was able to predict the learning functions in response to both a small and large fixed perturbation (light violet). **d)** The FIX and LINEAR models fail to predict the learning curve in the 3° condition. **e)** The hand angle ratio for the two error sizes. Early in learning, the ratio is large and then converges to a value slightly larger than 1. The best fitting model assumes that the labile and stable processes are modulated in a common manner, a signature of a system in which one signal gates the other. Shaded area and error bar indicate standard error.

In a second test of the gating hypothesis, we manipulated the duration of the inter-trial interval. The trial-by-trial change in hand angle arising from the labile process should decrease with the passage of time (Fig 6a). If this process gates the learning rate of the stable process, increasing the ITI should also decrease the

433 learning rate of the stable process (Fig 6b left). Alternatively, if the two processes operate in parallel
434 (PARALLEL model), the operation of the stable process should not be influenced by variation in ITI (Fig 6b
435 right).
436
437 To evaluate these predictions, we first used a trial-by-trial design (Exp 9) to measure the change in hand
438 angle when the ITI was short (0 s) or long (7 s). As predicted the change was attenuated in the 7 s condition
439 (Fig 6a). We then used a block design with either the short or long ITI (Exp 10). We ignored the first 5 trials
440 because these would still have a contribution from the labile process. Focusing on trials 6-10, we found
441 that the learning rate was higher in the short ITI condition, consistent with the prediction of the CPC model
442 (Fig 6d). Moreover, the long ITI condition showed a smaller retention in early washout but not in late
443 washout (Fig 6e), supporting the idea of the two learning processes.
444
445 However, the CPC model fails to capture one prominent feature in these data, the convergence of the two
446 functions at asymptote (Fig 6c & f). The basic CPC predicts that the advantage for the short ITI condition
447 should hold for the entire experiment, resulting in a lower asymptote for the long ITI condition. To address
448 this discrepancy, we modified the CPC model (Fig 6g), adding a recurrent suppression pathway from the
449 DCN to the inferior olive[65,66] (RSCPC). This inhibitory signal will reduce the strength of the error signal to
450 the PCs[67]. If we assume that the strength of this suppression decays across the ITI, the RSCPC generates
451 learning functions that provide good fits in both ITI conditions (Fig 6h & i, Fig S7). Importantly, after re-
452 estimating all parameters using this variant of the CPC, we observed negligible effects on the predictions
453 reported for the other experiments (Fig S8).

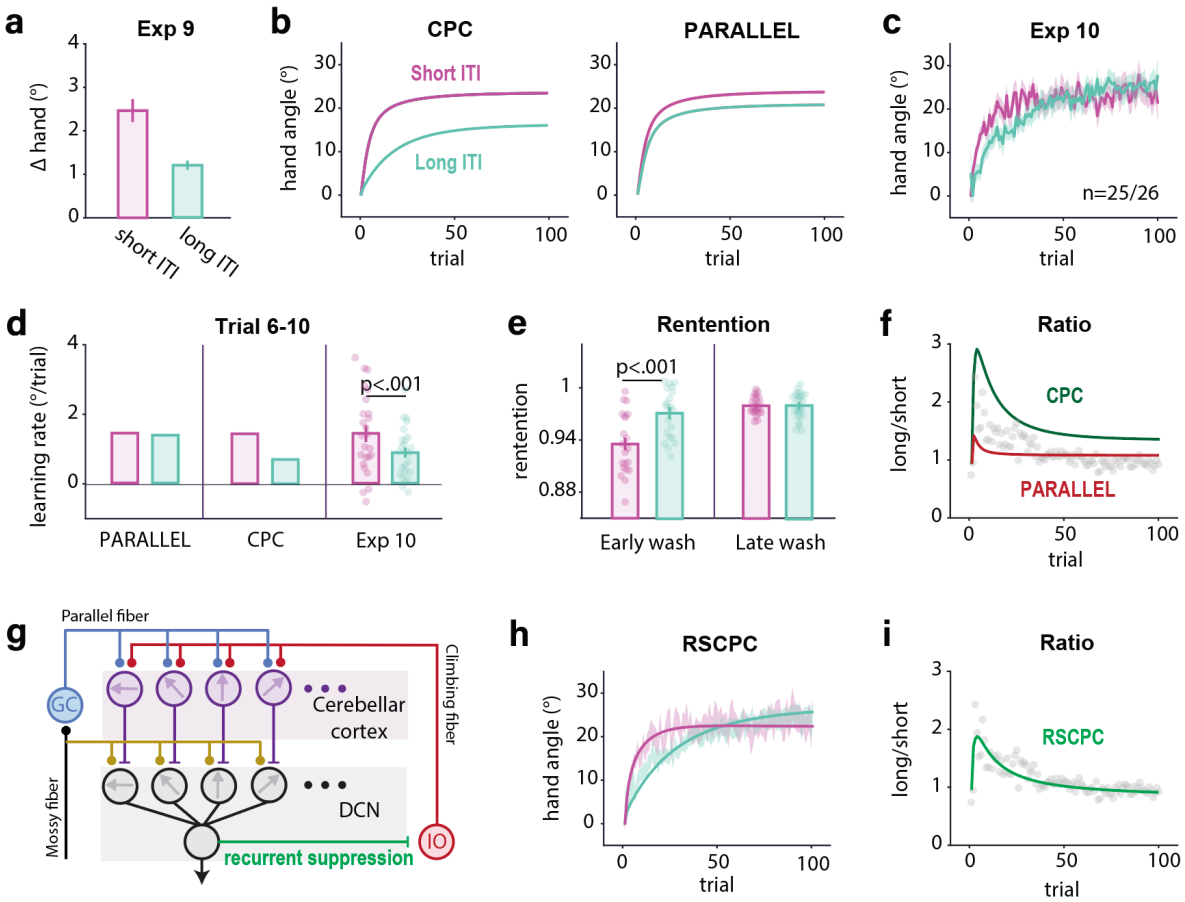


Fig. 6 The learning rate of the stable process is gated by the labile process. a) Trial-by-trial change in response to a variable perturbation with a short (Exp 5, p=0.5) or long ITI (Exp 9). **b)** Predictions of learning functions under the gating assumption of the CPC model and alternative model in which the two processes operate independently (PARALLEL). **c)** Learning functions in Exp 10 using either a short or long ITI. Consistent with the CPC model, the difference between the two functions is reduced over time. **d)** Model predictions and results from Exp 10 for the learning rate for trials 5-10. The learning rate is higher in the short ITI condition. **e)** The retention rate is larger in the initial no-feedback trials in the long ITI condition since the labile process is weakened by the passage of time. However, the retention rate is similar across the two ITI conditions later during washout, consistent with the hypothesis that only the stable process remains operative. **f)** Hand angle ratio between short ITI and long ITI condition deviates from predictions of both models. The ratio falls between the two model predictions early in training and is smaller than predicted by both models late in training. **g)** Modified RSCPC model includes inhibitory projection from DCN to the IO. This suppresses the error signal conveyed by the climbing fibers. This suppression is assumed to decay with time, becoming negligible in the long ITI condition in the RSCPC model. **h-i)** Predictions of the RSCPC model provide a good fit to the learning

curve (h) as well as the change in the ratio between the long and short ITI conditions (i) in Exp 10. Shaded area and error bar indicate standard error.

Discussion

To support flexible behavior, an organism needs to choose an action appropriate for a given context and execute a movement to achieve the desired outcome. Reaching this goal entails the operation of multiple learning mechanisms. A large body of work has sought to delineate the principles of these learning processes, with one prominent question centering on how the processes incorporate context and respond to uncertainty. With respect to the former, context not only helps specify the optimal action in a particular setting, but makes learning more efficient, providing a basis for both generalization and discrimination[55,68]. With respect to the latter, uncertainty, be it in terms of the environment or agent, has been shown to modulate the rate of learning[69,70]. A core challenge for models of learning requires specifying how context and uncertainty impact each learning mechanism.

Here we address this question with respect to the cerebellum, a subcortical structure long recognized as essential for certain types of error-based learning. We focused here on adaptation, the process by which error information is used to keep the sensorimotor system precisely calibrated in the face of fluctuations in the environment or state of the agent. To explore the impact of context and uncertainty on this learning process, we developed a population-coding model of the cerebellum incorporating two key features: 1) Each unit throughout the system is tuned to both movement direction and error direction, and 2) learning occurs at different rates in the cerebellar cortex and deep cerebellar nuclei, with the former characterized by a fast, labile process and the latter characterized by a slower, stable process. Our cerebellar population coding (CPC) model provides a parsimonious account of a diverse range of phenomena typically consider signatures of context-dependent learning as well as the impact of error consistency. The two-layer

493 structure offers new insight into the temporal dynamics of learning. (Table S1 summarizes comparisons
494 between the CPC models and other models in the field.)
495
496 Context Dependency as an Emergent Property of Population Coding
497 The CPC model offers a novel explanation for a well-known contextual effect, anterograde interference.
498 Classic models of this phenomenon focus on how the acquisition of new information is disrupted by the
499 reactivation of previously learned information[71]; as such, anterograde interference arises from
500 competition between different representations that are activated due to their contextual overlap. In
501 contrast, anterograde interference is an emergent property in the CPC model. Due to the different tuning
502 properties of neurons in the cerebellar cortex and DCN, the persistent activation of neurons to an error in
503 one direction will interfere with the response to an error in a different direction.
504
505 Importantly, there is no explicit role of context in the CPC model in the sense that a context does not
506 trigger the retrieval of its associated response. In this way, the CPC model diverges from classic models in
507 the behavior that emerges when a previously encountered context is re-experienced. Under such
508 circumstances, classic models predict savings in relearning given that the context facilitates the retrieval
509 of the appropriate response. Indeed, this is a key feature of COIN, a model of sensorimotor learning in
510 which context provides the critical retrieval cue[29]. In contrast, the CPC model accounts for the fact that
511 when a previously experienced perturbation is encountered, implicit adaptation not only fails to show
512 savings, but actually can show attenuation[46]. This attenuation in relearning is simply another
513 manifestation of anterograde interference: The opposite error is experienced during the washout phase
514 following initial learning and persistent changes in the state of the system interfere with learning when
515 the perturbation is reintroduced. Consistent with this hypothesis, no attenuation is observed in relearning

516 if the feedback is withheld during an extended washout phase before the second training (Guy
517 unpublished).
518
519 In a similar manner, we also find that implicit adaptation is insensitive to environmental uncertainty. *A*
520 *priori*, it would seem optimal for a learning system to be responsive to environmental uncertainty[69],
521 increasing the learning rate when faced with high volatility and decreasing it in the presence of a noisy or
522 highly stochastic environment. Such modulation has been observed in many studies of learning including
523 fear conditioning and some reward-based learning[70,72–75]. However, as we have shown in a previous
524 study[13], implicit adaptation is insensitive to perturbation variance and, in a related manner, insensitive to
525 the rate of switching between clockwise and counterclockwise perturbations (Exp 5).
526
527 We note that, with respect to the latter, the rate of learning is attenuated when the sign of the
528 perturbation changes relative to a fixed perturbation. But this attenuation is another manifestation of
529 anterograde interference; the degree of attenuation is not affected by the frequency of these sign changes.
530 In sum, the signatures of context-dependent learning and environmental uncertainty emerge naturally
531 from a population of tuned elements that operate in an inflexible manner.
532
533 Given the impressive flexibility in human motor learning, it might be surprising that implicit adaptation
534 does not explicitly track the context or uncertainty of the environment given the relevance of these two
535 factors for other learning systems[29,55,68]. We propose that this rigidity reflects a degree of modularity
536 between processes associated with action selection and those related to movement implementation. The
537 cerebellum is part of a system designed to use error information to ensure the accurate execution of a
538 desired movement. The emphasis here is on "desired movement" rather than "planned action" to
539 underscore the point that this system appears to operate independent of the task goal; indeed,

540 participants will adapt to sensory prediction errors even when the change in behavior is detrimental to
541 task success[5,8]. This modularity provides a means to keep the system properly calibrated across changes
542 in the internal state of the organism (e.g., perceptual biases, fatigue), factors that need not require a
543 change in action planning. In contrast, other learning systems are designed to use error information
544 related to task success to determine if the selected action was optimal given the current context. These
545 systems would be optimized to track contextual shifts in determining the appropriate policy. Consistent
546 with this hypothesis, contextual effects such as savings and sensitivity to uncertainty are observed in
547 adaptation tasks that benefit from changes in action selection[57,76,77].
548
549 Hierarchical Organization Within the Cerebellum for Implicit Adaptation
550 By using trial-by-trial and block designs, we verified that implicit adaptation operates on multiple
551 timescales. This idea was first articulated by Smith and colleagues[18] who described the parallel operation
552 of fast and slow processes in response to a large perturbation. Subsequent studies have led some
553 researchers to postulate that the fast and slow systems map onto explicit and implicit learning processes[49].
554 The current results provide new evidence that learning limited to the implicit system operates at different
555 timescales. However, rather than view these timescales as properties of learning processes that operate
556 in parallel (e.g., fast vs slow), our empirical and modeling results highlight a hierarchical organization in
557 which accumulated learning from a labile process will constrain the learning rate of a stable process. This
558 organization readily maps onto a two-layered network formed by cerebellar cortex and DCN, with the
559 output from the former gating learning within the latter.[42] Reflective of the hierarchical organization,
560 there is an asymmetric dependency such that the synaptic strength in the cerebellar cortex determines
561 PC output which will modulate the learning rate within the DCN.
562

563 The two-layer model provides an alternative explanation for another type of context-dependent learning,
564 contextual interference. The term is a bit of a misnomer since the phenomenon refers to the fact that,
565 while performance gains when training in multiple contexts is slower compared to training in a single
566 context, retention is better in the former[78,79]. As such, the exposure to multiple contexts during training
567 actually enhances learning as measured by long-term gains. Interestingly, this phenomenon is not limited
568 to skill acquisition tasks but is also observed in studies of implicit adaptation[80] (see Fig S9). As shown in
569 our simulations and Exp 10, contextual interference, at least in the context of implicit adaptation, is an
570 emergent property of the parallel operation of labile and stable learning processes. With multiple targets
571 (constituting multiple contexts), the rate of acquisition is slower compared to a single target since learning
572 from the labile process decays between successive reaches to a given target. However, early retention is
573 higher since the contribution of the labile process is small. Thus, as with anterograde interference,
574 contextual interference arises from the dynamics of the system without postulating any representation
575 of context.
576
577 Future Directions
578 We recognize that there are certain limitations with the CPC model in its current form. A key feature of
579 the model is the hierarchical organization of a two-layer network, one that we have attributed to the
580 organization of the cerebellar cortex and DCN. We have assumed that learning operates at different
581 timescales within these two layers, with plasticity operating at a faster time scale in the cerebellar cortex
582 compared to the DCN. The neurophysiological evidence is consistent with this assumption: While the
583 change of SS activation in the PCs can happen across a few trials[22,63], the change of dynamics in DCN can
584 be associated to behavior changes across sections or across days[25,38]. Nonetheless, this assumption should
585 be put to more direct evaluation. For example, experiments that examine the neural correlates of the rate
586 in the change in behavior. This could be accomplished by simultaneous recordings in the PCs and DCN.

587 Alternatively, it may be possible to look for anatomical-behavioral correlations in patients who vary in the
588 relative degree of atrophy in the cerebellar cortex or DCN.
589
590 A two-layered model is clearly a simplification. Indeed, to explain the asymptotic convergence in the long
591 and short ITI conditions, we had to incorporate a third layer into the model, creating a closed loop by
592 adding a projection from the DCN to the IO. While the anatomy supports the existence of this pathway,
593 to achieve convergence, we added two specific features to the dynamics of this pathway. First, inhibition
594 from the DCN to the IO exhibits a unique temporal dynamic, with its intensity decreasing over time.
595 Second, the projection is generic, inhibiting IO units independent of the directional tuning of the DCN
596 neuron. These two assumptions need to be tested in future physiological studies.
597
598 The current model does not address one prominent feature of cerebellar-dependent learning, namely the
599 sensitivity of this system to temporal regularities and to optimize timing of a learned response. The
600 adaptation phenomena modeled in the current paper do not entail a temporal component: Each trial
601 results in a directional error that is used to adjust the output of the system. However, timing is central to
602 other types of cerebellar-dependent learning such as eyeblink conditioning where the animal learns to
603 produce a conditioned response at an optimal time[81–85]. In models of eyeblink conditioning, the
604 interaction between granule cell, interneurons, and Purkinje cell activity will, across the population,
605 produce a sustained representation of the conditioned stimulus. The timing of the output will be shaped
606 by the interaction of this pattern with the input provided by the unconditioned stimulus. Generalizing the
607 CPC model to other forms of cerebellar-dependent learning will likely require adding these other elements
608 and dynamics.
609

610 The above sketches how population coding could be exploited to provide a more general account of
611 cerebellar function. We should also consider how the principles of population coding elucidated in the
612 CPC model might apply beyond the cerebellum. In particular, population-level dynamics might be
613 applicable to understand contextual effects in other domains such as perceptual learning[86,87].
614
615 **Methods**
616 Cerebellar Population Coding (CPC) model
617 The core features of the CPC model have been presented in the Results section. In the following section
618 we describe how the parameters of the model are determined.
619
620 We used an empirical approach to estimate the learning and forgetting rate for PF-PC synapses, using the
621 data from Exp 5 in which +/- 30° clamps were presented with a 50% switching probability. To measure
622 single trial learning, we calculate the change of hand angle between trial n and trial n-1, flipping the sign
623 when the clamp on trial n-1 was negative. To measure single trial forgetting, we calculate the change of
624 hand angle between trial n and trial n-1, flipping the sign when the clamp on trial n-2 was negative.
625
626 PF-PC forgetting (f) is the ratio of single-trial forgetting and single-trial-learning. By definition, retention
627 rate is 1- (f). We applied the same method to measure the retention for all trial-by-trial designs and this
628 gave us an f around 0.5. Model simulations indicate that this method can precisely estimate retention
629 when the perturbation is random. In all of these analyses, we excluded the first 50 trials since learning at
630 this early stage is influenced by both PC and DCN. For comparing the learning rate between early and late
631 training in a trial-by-trial design, we employed the same general approach but limited the analysis to the
632 first 50 trials to estimate early learning (Fig S5).
633

634 The baseline and maximal strength of MF-DCN synapses can be set to arbitrary values: We used 1 and
635 1.85 for m_{o} and $m_{\max}$, respectively. We measured the retention rate of the MF-DCN synapse (α)
636 empirically using the data from the no-feedback washout phase in Exp 1:

637 $$\alpha = \sqrt[10]{mean(\frac{y^{n+10}}{y^{n}})}$$

638 where y^n is the hand angle in trial n. The first 20 trials in the washout phase were excluded since they
639 may be contaminated by a labile process.

640

641 The learning rate of the PC (l) and DCN (ß) and the scaling factors (γ, ε) were jointly fitted from the
642 learning block in Exp 1 and the single trial learning in Exp5. This results in a parameter set as follow: l
643 = .05, f = .018, ß = 2, α = .5, γ= 0.15, ε = 130. These parameters were fixed in the simulations of all the
644 other experiments. The two exceptions are mentioned below.

645

646 First, in Exp 7, we examined how error size modulated learning using a block design in which the clamp
647 was at 3° or 30° clamp in separate conditions (between-subject). The prediction for the 30° clamp
648 condition was generated based on the parameter set described above. For the 3° clamp, we applied a
649 scale factor of 0.33 on CS activation:

650 $$[6]\ cs(3°) = 0.33 * cs(30°)$$

651 This value was based on the empirically observed values for 3° and 30° clamps in the trial-by-trial design
652 of Exp 6. Second, in Exp 10, we set the PF-PC retention rate for the long ITI conditions (f') to be 0.3, based
653 on the empirically-observed value in the trial-by-trial design of Exp 9.

654

655 <u>Recurrent Suppression Cerebellar Population Coding (RSCPC) model</u>

The results of Exp 9 led us to develop a post-hoc variant in which the output of the cerebellum modulates the input, an idea that is consistent with cerebellar anatomy and physiology[65,66]. The basic version of the CPC model predicts that learning in a long ITI condition will reach a lower asymptote compared to a short ITI condition. This occurs because the contribution of the labile process is suppressed in the long ITI condition. However, the results of Exp 10 showed that, with a sufficient number of trials, learning in the long ITI condition eventually reaches the same asymptote as in the short ITI condition. This observation led us to revise the model by adding an inhibitory pathway from the DCN to the inferior olive[65,66]., what we will refer to as the cerebellar population coding model with recurrent suppression (RSCPC).

We assume that the output of the DCN integrates the activation of directionally tuned units and that this signal serves as a generic inhibitory signal to the inferior olive. We implemented this recurrent suppression by subtracting a common value from the activation of cells tuned to all error directions in the inferior olive (IO):

$$[8] IO_i = 1 - \omega * \sum_i dDCN_i^n$$

$$[9] \text{ if } IO_i > 0: cs'_i = IO_i * cs_i;$$

$$\text{otherwise: } cs'_i = 0$$

where ω represents the strength of suppression. Given the assumption that ω decreases across time, we used separate parameter values of ω for the long and short ITI conditions $\sum_i dDCN_i^n$ is the sum of the change of all NCD units relative to their baseline activities. cs'_i is the corrected CS activation value after taking the DCN-IO pathway into the consideration and replaces the cs_i term in EQ [1-5]. The retention rates of the labile and stable processes (f, α) in the RSCPC model were set as in the basic two-layer model. The other parameters (l, ß, ε, ω) were jointly fitted from two data sets, the learning block in Exp 1 and the trial-by-trial condition in Exp 9. The parameter set is as follow: l = .1, f = .018, ß = 2, α = .5, γ = .2, ε = 210, $\omega(short)$ = 2.5, $\omega(long)$ = 0.

680

681 Alternative Models for Comparison

682 *Variants of the CPC Model*

683 To help clarify the importance of a two-layer model, we describe two variants of the CPC model. First, we

684 implemented a single-layer version of the CPC model by modifying Eq 4 to:

685 $$[7]\ DCN_i^{n+1} = m_i^{n+1}$$

686 In this version, the output of the system is solely determined by the strength of the MF-DCN.

687

688 Second, we implemented a model in which the labile and the stable processes operate in parallel

689 (PARALLEL) rather than hierarchical as in the CPC (and RSCPC) model. Since the stable process is insensitive

690 to ITI, we estimate the MF-DCN synapse (m) by simulations using a short ITI. The simulated value was then

691 used in simulations of the long ITI condition. For the labile process, the strength of the PF-PC synapse (w)

692 was measured separately for the two ITI conditions.

693

694 *State-space model*

695 We employed a standard version of a state-space model[18,48]:

696 $$[10]\ x(n+1) = a * x(n) + b(e,n)e(n) + \varepsilon_x(n)$$

697 where x is the internal estimate of the motor state (i.e., the hand movement required to compensate for

698 the perturbation), a is the retention factor, $e(n)$ is the size of the perturbation in trial n, b is the error

699 sensitivity for a given error size, and ε_x represents planning noise.

700

701 The actual motor response on trial n is given as:

702 $$[11]\ y(n) = x(n) + \varepsilon_y(n)$$

703 where y is the reaching direction relative to the target, determined by $x(n)$ and motor noise, ε_y .

Memory-of-Error Model (MoE)

The Memory-of-Error model describes how the learning rate in the state space model is modulated by experience. In the MoE model, error sensitivity (*b*) is set to an initial value that is modulated by errors that are experienced during training. Specifically, *b(e,n)* will increase if the error on trial n+1 shares the same sign and *b(e, n)* will decrease if the error on trial n+1 is of the opposite sign. This is formalized as:

$$[12]\ b(e(n), n+1) = \alpha * (b(e(n), n+1) - b0) + b0 + \beta * sign(e(n) * e(n+1))$$

where β and α are the learning rate and retention rate of *b*, respectively. Since the error size is fixed at 30° in our experiments, we replace $b(e)$ with a single value b.

Contextual Inference (COIN) model

We simulated the Contextual Interference (COIN) using the code provided by Heald et al.[29], focusing on its prediction with respect to savings and spontaneous recovery. We assumed that the introduction of a perturbation (e.g., clamped feedback) defines a new context and, as such, leads to the establishment of a new motor memory. Similarly, reversing the sign of the perturbation would define another context and thus require establishment of another memory. We simulated the clamps as if they were contingent rotations so that the learning can reach an asymptote. Before each movement, the output is determined by averaging the state of different contexts weighted by the expected probabilities of the contexts. Participants observed an error after each movement and update the state estimation.

<u>Behavioral Experiments</u>

Participants

A total of 451 participants (297 female, mean age = 28.0, SD = 5.3) were recruited through the website prolific.co. After eliminating participants who failed to meet our performance criteria (2.8%, see below), the analyses were based on data from 438 participants. Based on a survey included in a prescreening questionnaire, the participants were right-handed with normal or corrected-to-normal vision. The participants were paid based on a rate of $8/h. The protocol was approved by the Institutional Review Board at the University of California, Berkeley. Informed consent was obtained from all participants.

Apparatus

All of the behavioral experiments were conducted online using a web-based experimental platform, OnPoint[62], which is written in JavaScript and presented via Google Chrome. It is designed to operate on any laptop computer. Visual stimuli were presented on the laptop monitor and movements were produced on the computer trackpad. Data were collected and stored using Google Firebase.

Clamp rotation task

We applied clamp feedback in the experiments, under the assumption that learning in response to this type of feedback is limited to implicit, cerebellar-dependent sensorimotor recalibration. To start each trial, the participant moved the cursor to a white start circle (radius: 1% of the screen height) positioned in the center of the screen. After 500ms, the target, a blue circle (radius: 1% of the screen height) appeared with the radial distance set to 40% of the screen size. The target appears at -45°, a workspace location selected because is exhibits minimal bias across participants[88]. The participant was instructed to produce a rapid, out-and-back movement, attempting to intersect the target. If the movement time (from onset to time at which movement amplitude reached the target) was longer than 500ms, the message 'Too Slow' was presented on the screen for 500ms.

There were three types of feedback. On veridical feedback trials, the position of the cursor moved was matched to the position of the hand, subject to the translation in reference frames (screen assumed to be vertical, hand movement assumed to be horizontal) and scaling (trackpad space expanded to encompass most of the screen). On clamped feedback trials, the cursor followed a fixed path. As with veridical feedback, the radial location of the cursor was based on the radial extent of the participant's hand. However, the angular position of the cursor was independent of the position of the hand, instead determined relative to the position of the target. The clamp angle was set at 30° relative to the target except for Exp 6 and 8 (see below). On no feedback trials, the cursor was blanked at movement onset.

On veridical and clamped feedback trials, after the amplitude of the movement reached the target distance, the cursor was presented at the target distance for another 50ms then it disappeared. Target disappeared after 200ms. The cursor was then reset to a random position on an invisible circle with a radius equal to 10% of the target distance and the participant moved the cursor back to the start circle.

At the onset of the first block of trials involving perturbed feedback, the experiment was paused and a set of instructions were presented to describe the clamped feedback. The participant was informed that the cursor would no longer be linked to their movement but rather would follow a fixed path on all trials. The participant was instructed to always reach directly to the target, ignoring the cursor. These instructions were then repeated twice to emphasize the atypical nature of the feedback. After the first 10 trials with clamped feedback, a new instruction screen appeared in which the participant was asked to indicate where they were aiming on each trial. If the participant indicated they were reaching somewhere other than the target, the experiment was terminated.

Each experiment started with two baseline blocks: First a no-feedback block of 10 trials and second, a veridical feedback block of 10 trials. For experiments using a block design (direction and size of perturbation remains constant), the direction of the clamp (counterclockwise, CCW; clockwise; CW) was counterbalanced across participants.

Experiment 1

Exp 1 was designed to determine the parameters of the CPC model. There was a total of 180 trials. The two baseline blocks were followed by a learning block of 100 trials with clamped feedback with learning expected to reach an asymptotic level in response to a fixed perturbation. This was followed by a final no-feedback block of 60 trials. 30 participants were recruited for Exp 1 (29 valid, 5 males, age: 27.4 ± 4.9 years).

Experiment 2

Exp 2 was designed to evaluate different models of asymptotic adaptation. The 10-trial feedback baseline were followed by a learning block of 100 trials with clamped feedback. Then the last 60 trials alternated between no-feedback and clamp feedback trials (half-wash phase). 40 participants were recruited for Exp 2 (38 valid, 8 males, age: 30.7 ± 6.6 years).

Experiment 3

Exp 3 was designed to measure antegrade interference. The baseline and initial perturbation blocks were as in Exp 2. For the final block (150 trials), the direction of the clamp was reversed (e.g., from 30° to -30°). 30 participants were recruited for Exp 3 (30 valid, 10 males, age: 30.3 ± 4.3 years).

Experiment 4

Exp 4 was designed to assess spontaneous recovery and savings in implicit adaptation. The baseline and initial perturbation blocks were as in Exp 2. We then included a 15-trial block with the clamp reversed under the assumption that this would be a sufficient number of trials to bring the hand angle back to baseline. This was followed by no-feedback block (35 trials) to examine spontaneous recovery and then a 100-trial relearning block in which the clamp feedback was identical to that used in the first perturbation block. 34 participants were recruited for Exp 4 (34 valid, 16 males, age: 22.7 ± 4.8 years).

Experiment 5

Exp 5 examined how the consistency of the perturbation influenced implicit adaptation. The first blocks were identical to Exp 4, providing initial exposure to clamped feedback and then a reversed clamp to bring the hand angle back to baseline. This was followed by a 300-trial block in which the clamp changed sign in a probabilistic manner. The probability of a sign change was either 90%, 50%, and 12.5% in a between-subject manipulation. The sequence of clamps was preset to ensure that clockwise and counterclockwise occurred on 50% of the trials each across the 300 trials. The experiment ended with a relearning block in which the initial perturbation was presented for 100 trials. 36/40/36 participants were recruited for 90%, 50%, and 12.5% conditions respectively (34/38/33 valid, 37 males, age: 28.6 ± 5.5 years).

Experiment 6

To estimate the learning rate and retention at top layer of the CPC model, the PF-PC synapse, we employed a trial-by-trial design in which the error size and direction varied across trials. After the two baseline sections, participants completed a 540-trial random perturbation block. Here the clamp size ranged from -135° to 135° in steps of 1°. The size/direction was determined at random with the constraint that each clamp was selected once every 270 trials. 72 participants were recruited for Exp 6 (70 valid, 25 males, age: 26.2 ± 5.2 years).

Experiment 7

Experiment 7 was designed to measure the time course of retention during the initial washout phase. After the two baseline sections, The perturbation block consisted of 31 mini-blocks, each composed of 10 trials with clamped feedback and 10 trials without feedback (620 trials). 57 participants were recruited for Exp 7 (57 valid, 12 males, age: 28.3 ± 5.4 years).

Experiment 8

Experiment 8 examined how error size influences learning in a block design. Two groups of participants experience a 10-trial feedback baseline and a 100-trial perturbation block in which the clamp size was either 3° or 30°. 36 participants were recruited for Exp 8 (35 valid, 18 males, age: 30.8 ± 7.7 years).

Experiment 9

To quantify the temporal dynamics of labile processes, we performed a trial-by-trial design with extended inter-trial intervals (ITI) in Exp 9. For the long ITI, the interval between the end of one trial and the start of the next trial was 6s, 7s, or 8s, randomized across trials. The message "wait" was displayed on the monitor after each trial. Exp 9 included two baseline blocks and a 180-trial learning block in which a 30° perturbation was randomly selected to be either clockwise or counterclockwise, subject to the constraint that each direction occurred four times every 8 trials. For the short ITI condition, we used the data from Exp 5 for the trial-by-trial condition (0 s ITI). 28 participants were recruited for each condition (27 valid, 13 males, age: 28.1 ± 4.8 years).

Experiment 10

To understand how the labile and stable learning processes are jointly modulated by time, we perform a block design in Exp 10. We followed a similar design to Exp 1, with only one notable modification. We included a 10-trial filmization block following the two baseline blocks to demonstrate the clamp feedback. The clamp size in the filmization block varied from -90° to 90° across trials to show that the cursor is unaffected by the direction of hand movement. To avoid the influence of pre-exposure to the error signal on learning, the filmization block utilized a different target (45°) from the other blocks (315°). Two groups of participants perform the task with either long ITI (6-8s) or short ITI (0s). 26 participants were recruited for each condition (51 valid, 21 males, age: 26.8 ± 4.6 years).

Data analyses

Hand angle was calculated as the angle difference a line from the start position to the target and a line from the start position to the hand position at the target radius. Positive values indicate hand angles in the opposite direction of the perturbation, the direction one would expect due to adaptation. Trials with a movement duration longer than 500 ms or an error larger than 70° were excluded from the analyses. We excluded the entire data from participants who had less than 70% valid trials (2.8% participants). Between-condition comparisons were performed with t-tests or ANOVAs. Learning and relearning are compared by paired-t-test. For all the statistical tests, we confirmed that the data met the assumptions of a Gaussian distribution and homoscedasticity.

Author contributions

T.W., R.B.I. contributed to the conceptual development of this project. T.W. collected the data, analyzed the data, prepared the figures, and wrote the initial draft of the paper, with all the authors involved in the editing process.

Acknowledgment

This study is funded by the NIH (grants NS116883 and NS105839).

Competing interests

RI is a co-founder with equity in Magnetic Tides, Inc.

References

1. Ito, M. Mechanisms of motor learning in the cerebellum. *Brain Res.* **886**, 237–245 (2000).
2. Albus, J. S. A theory of cerebellar function. *Math. Biosci.* **10**, 25–61 (1971).
3. Kawato, M. Internal models for motor control and trajectory planning. *Curr. Opin. Neurobiol.* **9**, 718–727 (1999).
4. Ebner, T. J. & Pasalar, S. Cerebellum predicts the future motor state. *Cerebellum* **7**, 583–588 (2008).
5. Morehead, J. R., Taylor, J. A., Parvin, D. E. & Ivry, R. B. Characteristics of Implicit Sensorimotor Adaptation Revealed by Task-irrelevant Clamped Feedback. *J. Cogn. Neurosci.* **29**, 1061–1074 (2017).
6. Taylor, J. A., Klemfuss, N. M. & Ivry, R. B. An Explicit Strategy Prevails When the Cerebellum Fails to Compute Movement Errors. *Cerebellum* **9**, 580–586 (12/2010).
7. Taylor, J. A. & Ivry, R. B. Cerebellar and prefrontal cortex contributions to adaptation, strategies, and reinforcement learning. *Prog. Brain Res.* **210**, 217–253 (2014).
8. Mazzoni, P. & Krakauer, J. W. An implicit plan overrides an explicit strategy during visuomotor adaptation. *J. Neurosci.* **26**, 3642–3645 (2006).
9. Taylor, J. A., Krakauer, J. W. & Ivry, R. B. Explicit and Implicit Contributions to Learning in a Sensorimotor Adaptation Task. *Journal of Neuroscience* **34**, 3023–3032 (2014).
10. Wang, T. & Taylor, J. A. Implicit adaptation to mirror reversal is in the correct coordinate system but the wrong direction. *J. Neurophysiol.* **126**, 1478–1489 (2021).
11. Wilterson, S. A. & Taylor, J. A. Implicit Visuomotor Adaptation Remains Limited after Several Days of Training. *eNeuro* **8**, (2021).
12. Avraham, G., Keizman, M. & Shmuelof, L. Environmental consistency modulation of error sensitivity during motor adaptation is explicitly controlled. *J. Neurophysiol.* **123**, 57–69 (2020).

886 13. Wang, T., Avraham, G., Tsay, J. S., Abram, S. J. & Ivry, R. B. Perturbation Variability Does Not Influence
887 Implicit Sensorimotor Adaptation. *bioRxiv* 2023.01.27.525949 (2023)
888 doi:10.1101/2023.01.27.525949.
889 14. Albert, S. T. *et al.* Competition between parallel sensorimotor learning systems. *Elife* **11**, (2022).
890 15. Avraham, G., Taylor, J. A., Breska, A., Ivry, R. B. & McDougle, S. D. Contextual effects in sensorimotor
891 adaptation adhere to associative learning rules. *bioRxiv* (2020) doi:10.1101/2020.09.14.297143.
892 16. Denny-Brown, D. Conditioned reflexes: An investigation of the physiological activity of the cerebral
893 cortex. *Nature* **121**, 662–664 (1928).
894 17. Ebbinghaus, H. Memory: a contribution to experimental psychology. *Ann. Neurosci.* **20**, 155–156
895 (2013).
896 18. Smith, M. A., Ghazizadeh, A. & Shadmehr, R. Interacting Adaptive Processes with Different
897 Timescales Underlie Short-Term Motor Learning. *PLoS Biol.* **4**, e179 (2006).
898 19. Kawato, M. Feedback-Error-Learning Neural Network for Supervised Motor Learning. in *Advanced*
899 *Neural Computers* (ed. Eckmiller, R.) 365–372 (North-Holland, 1990).
900 20. Wolpert, D. M., Miall, R. C. & Kawato, M. Internal models in the cerebellum. *Trends Cogn. Sci.* **2**, 338–
901 347 (1998).
902 21. Apps, R. & Garwicz, M. Anatomical and physiological foundations of cerebellar information
903 processing. *Nat. Rev. Neurosci.* **6**, 297–311 (2005).
904 22. Herzfeld, D. J., Kojima, Y., Soetedjo, R. & Shadmehr, R. Encoding of action by the Purkinje cells of the
905 cerebellum. *Nature* **526**, 439–442 (2015).
906 23. Sedaghat-Nejad, E., Pi, J. S., Hage, P., Fakharian, M. A. & Shadmehr, R. Synchronous spiking of
907 cerebellar Purkinje cells during control of movements. *Proceedings of the National Academy of*
908 *Sciences* **119**, e2118954119 (2022).

909 24. Junker, M. *et al.* Learning from the past: A reverberation of past errors in the cerebellar climbing fiber
910 signal. *PLoS Biol.* **16**, e2004344 (2018).
911 25. Carulli, D. *et al.* Cerebellar plasticity and associative memories are controlled by perineuronal nets.
912 *Proc. Natl. Acad. Sci. U. S. A.* **117**, 6855–6865 (2020).
913 26. Raymond, J. L., Lisberger, S. G. & Mauk, M. D. The cerebellum: a neuronal learning machine? *Science*
914 **272**, 1126–1131 (1996).
915 27. Lisberger, S. G., Pavelko, T. A. & Broussard, D. M. Responses during eye movements of brain stem
916 neurons that receive monosynaptic inhibition from the flocculus and ventral paraflocculus in
917 monkeys. *J. Neurophysiol.* **72**, 909–927 (1994).
918 28. Lee, K. H. *et al.* Circuit mechanisms underlying motor memory formation in the cerebellum. *Neuron*
919 **86**, 529–540 (2015).
920 29. Heald, J. B., Lengyel, M. & Wolpert, D. M. Contextual inference underlies the learning of sensorimotor
921 repertoires. *Nature* **600**, 489–493 (2021).
922 30. Haruno, M., Wolpert, D. M. & Kawato, M. Hierarchical MOSAIC for movement generation. *Int. Congr.*
923 *Ser.* **1250**, 575–590 (2003).
924 31. Collins, A. & Koechlin, E. Reasoning, learning, and creativity: frontal lobe function and human
925 decision-making. *PLoS Biol.* **10**, e1001293 (2012).
926 32. Marr, D. A theory of cerebellar cortex. *J. Physiol.* **202**, 437–470 (1969).
927 33. Barash, S. *et al.* Saccadic dysmetria and adaptation after lesions of the cerebellar cortex. *J. Neurosci.*
928 **19**, 10931–10939 (1999).
929 34. Witter, L., Canto, C. B., Hoogland, T. M., de Gruijl, J. R. & De Zeeuw, C. I. Strength and timing of motor
930 responses mediated by rebound firing in the cerebellar nuclei after Purkinje cell activation. *Front.*
931 *Neural Circuits* **7**, 133 (2013).

932 35. Ito, M. & Kano, M. Long-lasting depression of parallel fiber-Purkinje cell transmission induced by
933 conjunctive stimulation of parallel fibers and climbing fibers in the cerebellar cortex. *Neurosci. Lett.*
934 **33**, 253–258 (1982).
935 36. Hutter, S. A. & Taylor, J. A. Relative sensitivity of explicit reaiming and implicit motor adaptation. *J.*
936 *Neurophysiol.* **120**, 2640–2648 (2018).
937 37. Wei, K. & Körding, K. Relevance of Error: What Drives Motor Adaptation? *J. Neurophysiol.* **101**, 655–
938 664 (02/2009).
939 38. Moscato, L. *et al.* Long-Lasting Response Changes in Deep Cerebellar Nuclei in vivo Correlate With
940 Low-Frequency Oscillations. *Front. Cell. Neurosci.* **13**, 84 (2019).
941 39. Ito, M. Cerebellar control of the vestibulo-ocular reflex--around the flocculus hypothesis. *Annu. Rev.*
942 *Neurosci.* **5**, 275–296 (1982).
943 40. Medina, J. F., Garcia, K. S., Nores, W. L., Taylor, N. M. & Mauk, M. D. Timing mechanisms in the
944 cerebellum: testing predictions of a large-scale computer simulation. *J. Neurosci.* **20**, 5516–5525
945 (2000).
946 41. Garcia, K. S. & Mauk, M. D. Pharmacological analysis of cerebellar contributions to the timing and
947 expression of conditioned eyelid responses. *Neuropharmacology* **37**, 471–480 (1998).
948 42. Herzfeld, D. J., Hall, N. J., Tringides, M. & Lisberger, S. G. Principles of operation of a cerebellar
949 learning circuit. *Elife* **9**, (2020).
950 43. Shadmehr, R. Population coding in the cerebellum: a machine learning perspective. *J. Neurophysiol.*
951 **124**, 2022–2051 (2020).
952 44. Kim, H. E., Avraham, G. & Ivry, R. B. The Psychology of Reaching: Action Selection, Movement
953 Implementation, and Sensorimotor Learning. *Annu. Rev. Psychol.* **72**, 61–95 (2021).
954 45. Avraham, G., Taylor, J. A., Breska, A., Ivry, R. B. & McDougle, S. D. Contextual effects in sensorimotor
955 adaptation adhere to associative learning rules. *Elife* **11**, e75801 (2022).

46. Avraham, G., Morehead, J. R., Kim, H. E. & Ivry, R. B. Reexposure to a sensorimotor perturbation produces opposite effects on explicit and implicit learning processes. *PLoS Biol.* **19**, e3001147 (2021).

47. Tsay, J. S., Parvin, D. E. & Ivry, R. B. Continuous reports of sensed hand position during sensorimotor adaptation. *J. Neurophysiol.* **124**, 1122–1130 (2020).

48. Shadmehr, R. & Mussa-Ivaldi, F. A. Adaptive representation of dynamics during learning of a motor task. *J. Neurosci.* **14**, 3208–3224 (1994).

49. McDougle, S. D., Bond, K. M. & Taylor, J. A. Explicit and Implicit Processes Constitute the Fast and Slow Processes of Sensorimotor Learning. *Journal of Neuroscience* **35**, 9568–9579 (七月 1 2015).

50. Morehead, J. R. & Smith, M. The magnitude of implicit sensorimotor adaptation is limited by continuous forgetting. *Abstract. Advances in Motor Learning & Motor Control* (2017).

51. Lerner, G. *et al.* The origins of anterograde interference in visuomotor adaptation. *Cereb. Cortex* **30**, 4000–4010 (2020).

52. Brashers-Krug, T., Shadmehr, R. & Bizzi, E. Consolidation in human motor memory. *Nature* **382**, 252–255 (1996).

53. Leow, L.-A., Hammond, G. & de Rugy, A. Anodal motor cortex stimulation paired with movement repetition increases anterograde interference but not savings. *Eur. J. Neurosci.* **40**, 3243–3252 (2014).

54. Sing, G. C. & Smith, M. A. Reduction in learning rates associated with anterograde interference results from interactions between different timescales in motor adaptation. *PLoS Comput. Biol.* **6**, e1000893 (2010).

55. Heald, J. B., Lengyel, M. & Wolpert, D. M. Contextual inference in learning and memory. *Trends Cogn. Sci.* (2022) doi:10.1016/j.tics.2022.10.004.

56. Ingram, J. N., Flanagan, J. R. & Wolpert, D. M. Context-dependent decay of motor memories during skill acquisition. *Curr. Biol.* **23**, 1107–1112 (2013).

979 57. Herzfeld, D. J., Vaswani, P. A., Marko, M. K. & Shadmehr, R. A memory of errors in sensorimotor
980 learning. *Science* **345**, 1349–1353 (2014).
981 58. Kording, K. P., Tenenbaum, J. B. & Shadmehr, R. The dynamics of memory as a consequence of
982 optimal adaptation to a changing body. *Nat. Neurosci.* **10**, 779–786 (2007).
983 59. Kim, H. E., Morehead, J. R., Parvin, D. E., Moazzezi, R. & Ivry, R. B. Invariant errors reveal limitations
984 in motor correction rather than constraints on error sensitivity. *Commun Biol* **1**, 19 (12/2018).
985 60. Tsay, J. S. *et al.* The effect of visual uncertainty on implicit motor adaptation. *J. Neurophysiol.* **125**,
986 12–22 (2021).
987 61. Wang, T., Avraham, G., Tsay, J. & Ivry, R. The Effect of Perturbation Variability on Sensorimotor
988 Adaptation Does Not Require an Implicit Memory of Errors. *bioRxiv* 2022.05.30.493844 (2022)
989 doi:10.1101/2022.05.30.493844.
990 62. Tsay, J. S., Ivry, R. B., Lee, A. & Avraham, G. Moving outside the lab: The viability of conducting
991 sensorimotor learning studies online. *Neurons, Behavior, Data analysis, and Theory* (2021)
992 doi:10.51628/001c.26985.
993 63. Medina, J. F. & Lisberger, S. G. Links from complex spikes to local plasticity and motor learning in the
994 cerebellum of awake-behaving monkeys. *Nat. Neurosci.* **11**, 1185–1192 (2008).
995 64. Herzfeld, D. J., Kojima, Y., Soetedjo, R. & Shadmehr, R. Encoding of error and learning to correct that
996 error by the Purkinje cells of the cerebellum. *Nat. Neurosci.* **21**, 736–743 (5/2018).
997 65. Best, A. R. & Regehr, W. G. Inhibitory regulation of electrically coupled neurons in the inferior olive
998 is mediated by asynchronous release of GABA. *Neuron* **62**, 555–565 (2009).
999 66. De Zeeuw, C. I. & Ruigrok, T. J. Olivary projecting neurons in the nucleus of Darkschewitsch in the cat
1000 receive excitatory monosynaptic input from the cerebellar nuclei. *Brain Res.* **653**, 345–350 (1994).
1001 67. Bengtsson, F. & Hesslow, G. Cerebellar control of the inferior olive. *Cerebellum* **5**, 7–14 (2006).

1002 68. Collins, A. G. E. & McDougle, S. D. Context is key for learning motor skills. *Nature* vol. 600 387–388
1003 (2021).
1004 69. Piray, P. & Daw, N. D. A model for learning based on the joint estimation of stochasticity and volatility.
1005 *Nat. Commun.* **12**, 6587 (2021).
1006 70. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of
1007 information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
1008 71. Wixted, J. T. The psychology and neuroscience of forgetting. *Annu. Rev. Psychol.* **55**, 235–269 (2004).
1009 72. Nassar, M. R. *et al.* Rational regulation of learning dynamics by pupil-linked arousal systems. *Nat.*
1010 *Neurosci.* **15**, 1040–1046 (2012).
1011 73. Iglesias, S. *et al.* Hierarchical prediction errors in midbrain and basal forebrain during sensory learning.
1012 *Neuron* **80**, 519–530 (2013).
1013 74. McGuire, J. T., Nassar, M. R., Gold, J. I. & Kable, J. W. Functionally dissociable influences on learning
1014 rate in a dynamic environment. *Neuron* **84**, 870–881 (2014).
1015 75. Soltani, A. & Izquierdo, A. Adaptive learning under expected and unexpected uncertainty. *Nat. Rev.*
1016 *Neurosci.* **20**, 635–644 (2019).
1017 76. Herzfeld, D. J. & Shadmehr, R. Cerebellum estimates the sensory state of the body. *Trends in cognitive*
1018 *sciences* vol. 18 66–67 (2014).
1019 77. Gonzalez Castro, L. N., Hadjiosif, A. M., Hemphill, M. A. & Smith, M. A. Environmental consistency
1020 determines the rate of motor adaptation. *Curr. Biol.* **24**, 1050–1061 (2014).
1021 78. Tsay, J. S., Irving, C. & Ivry, R. B. Signatures of contextual interference in implicit sensorimotor
1022 adaptation. *Proc. Biol. Sci.* **290**, 20222491 (2023).
1023 79. Collins, A. G. E. *The tortoise and the hare: interactions between reinforcement learning and working*
1024 *memory*. http://biorxiv.org/lookup/doi/10.1101/234724 (2017).

1025 80. Shea, J. B. & Morgan, R. L. Contextual interference effects on the acquisition, retention, and transfer
1026 of a motor skill. *J. Exp. Psychol. Hum. Learn.* **5**, 179–187 (1979).
1027 81. Johansson, F., Hesslow, G. & Medina, J. F. Mechanisms for motor timing in the cerebellar cortex. *Curr*
1028 *Opin Behav Sci* **8**, 53–59 (2016).
1029 82. Yeo, C. H., Hardiman, M. J. & Glickstein, M. Classical conditioning of the nictitating membrane
1030 response of the rabbit. I. Lesions of the cerebellar nuclei. *Exp. Brain Res.* **60**, 87–98 (1985).
1031 83. Frey, P. W. & Ross, L. E. Rabbit eyelid conditioning: Effects of age, interstimulus interval, and intertrial
1032 interval. *Dev. Psychobiol.* **1**, 276–279 (1968).
1033 84. Schneiderman, N. & Gormezano, I. Conditioning of the nictitating membrane of the rabbit as a
1034 function of cs-us interval. *J. Comp. Physiol. Psychol.* **57**, 188–195 (1964).
1035 85. Smith, M. C. CS-US interval and US intensity in classical conditioning of the rabbit's nictitating
1036 membrane response. *J. Comp. Physiol. Psychol.* **66**, 679–687 (1968).
1037 86. Gilbert, C. D., Li, W. & Piech, V. Perceptual learning and adult cortical plasticity. *J. Physiol.* **587**, 2743–
1038 2751 (2009).
1039 87. Schoups, A. A., Vogels, R. & Orban, G. A. Human perceptual learning in identifying the oblique
1040 orientation: retinotopy, orientation specificity and monocularity. *J. Physiol.* **483**, 797–810 (1995).
1041 88. Vindras, P., Desmurget, M., Prablanc, C. & Viviani, P. Pointing Errors Reflect Biases in the Perception
1042 of the InitialHand Position. *J. Neurophysiol.* **79**, 3290–3294 (1998).
1043 89. Tsay, J. S., Irving, C. & Ivry, R. B. Signatures of contextual interference in implicit sensorimotor
1044 adaptation. *bioRxiv* (2022) doi:10.1101/2022.07.03.498608.
1045

Supplementary Information

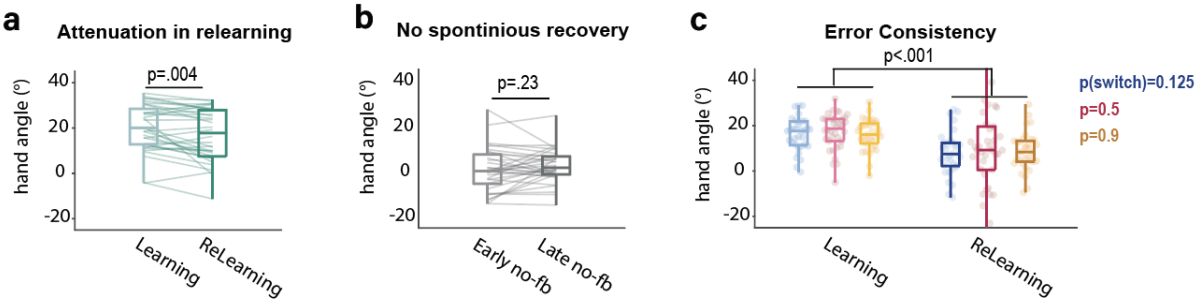


Fig. S1 Effect of experience and error consistency on implicit adaptation. a) Attenuation in relearning in Exp 4. Adaptation was attenuated in response to re-exposure to a perturbation compared to the initial exposure (t(33)=3.1, p=0.004) Data are averaged across each training phase. **b)** Spontaneous recovery was not observed in Exp 4 during the no-feedback phase after washout. Hand angle over the first 5 trials of the no-feedback phase (Early) is similar to hand angle over the last 5 trials (Late, t(33)=1.2, p=0.23). c) Error consistency did not affect adaptation during initial learning and during relearning in Exp 5. A mixed ANOVA showed a main effect of learning/relearning, (F(1,101)=37.7, p<0.001), similar to the antegrade interference observed in Exp 3. There was no effect of error consistency (F(2,101)=0.18, p=0.84) or interaction between phase and error consistency (F(2,101)=0.12, p=0.88). Box plots indicate median, max and min values, and 25% and 75% quartiles.

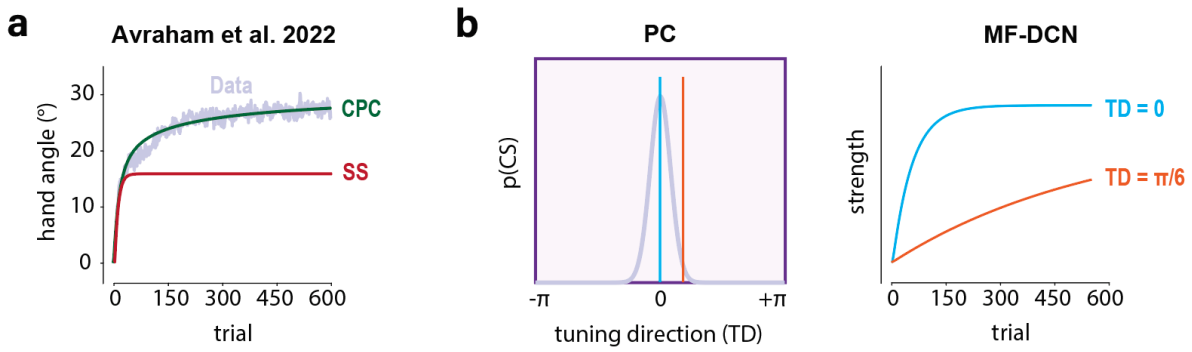


Fig. S2 The CPC model incorporates adaptation at different rates. a) Modeling adaptation learning functions frequently requires postulating multiple learning processes rather than a single-process state-space (SS) model. The panel depicts a learning curve from Avraham et al (2022) in which participants were exposed to a 30° (intermixed with 0° clamps, but not relevant for the current point). A single-process state-space model can capture the rapid change in hand angle early in learning but then saturates, failing to capture the gradual increase in late adaptation. The CPC model simultaneously captures early and late adaptation through the operation of multiple learning processes. b) Cells with a tuning direction (TD) aligned to the error direction (blue) respond strongly to the error (left), driving rapid early adaptation and saturate quickly (right). Cells with tuning misaligned with the error direction (orange, π/6 in this example) have a relatively low error response early in training but make a relatively large contribution late in training. Note that "fast" and "slow" emerge from the tuning properties of units within a single layer (DCN).

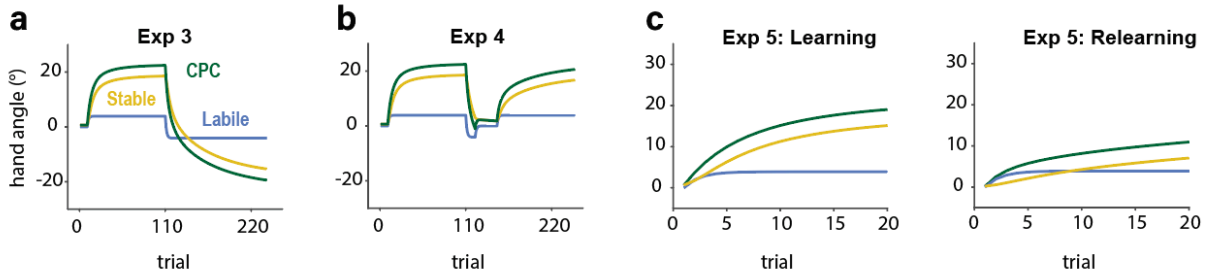


Fig. S3 Predicted time course of stable and labile processes in Exp 3-5. The stable process is responsible for anterograde interference (a) and attenuation in relearning (b-c). The labile process does not make a significant contribution to either phenomenon because of its low retention rate.

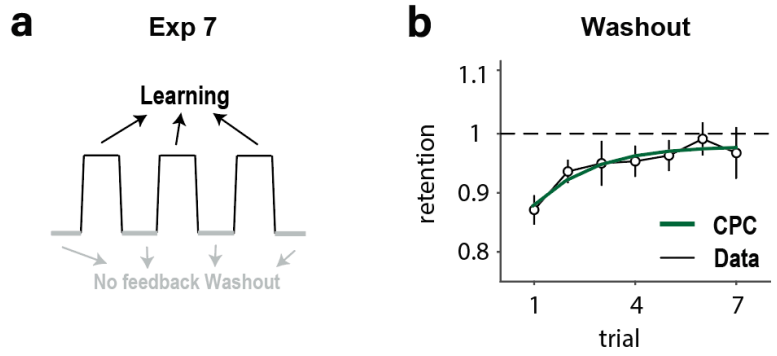
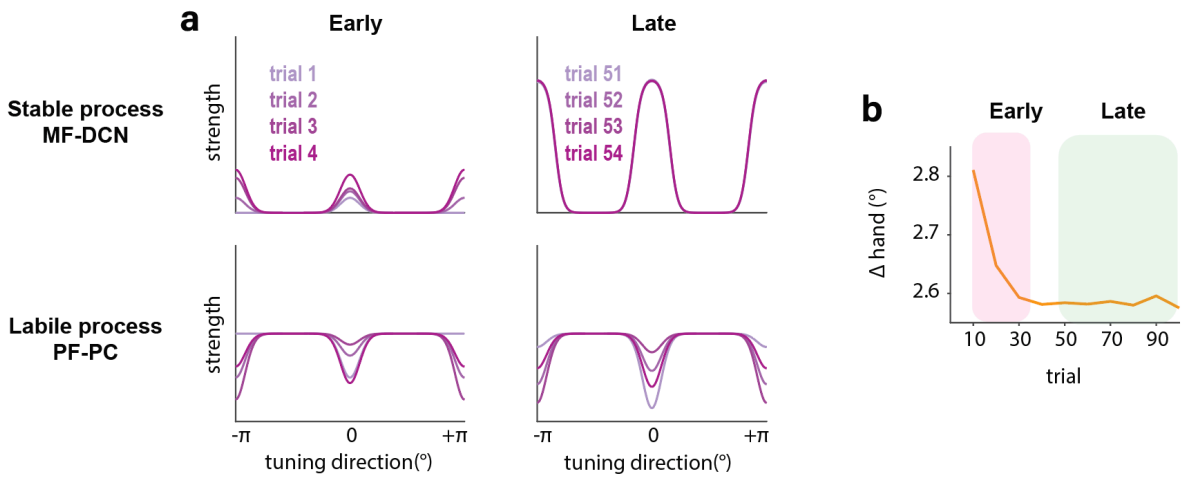


Fig. S4 Retention increases during the initial washout trials. a) To provide a stronger test of how the rate of retention changes (Exp 1), Exp 7 included mini-blocks (10 trials/mini-block) that alternated between clamp and no feedback trials. **B**) We estimated the change in retention rate over time by averaging by trial number across the no feedback blocks. Retention is relatively low in the first trials of the washout block and gradually rises ($F(6,264)=4.64$, $p<0.001$). The dark green curve shows the fit of the CPC model.



1087

1088 **Fig. S5 Contribution of stable and labile processes in response to variable perturbations. a)** The stable

1089 process (top) contributes to learning during early training and has saturated by the 50th trial. The

1090 contribution of the labile process (bottom) remains similar throughout training. **b)** Change in hand angle

1091 as a function of trial number when the size and direction of the perturbation varies across trials. The

1092 change of hand angle is larger in early training because the stable process has not saturated.

1093

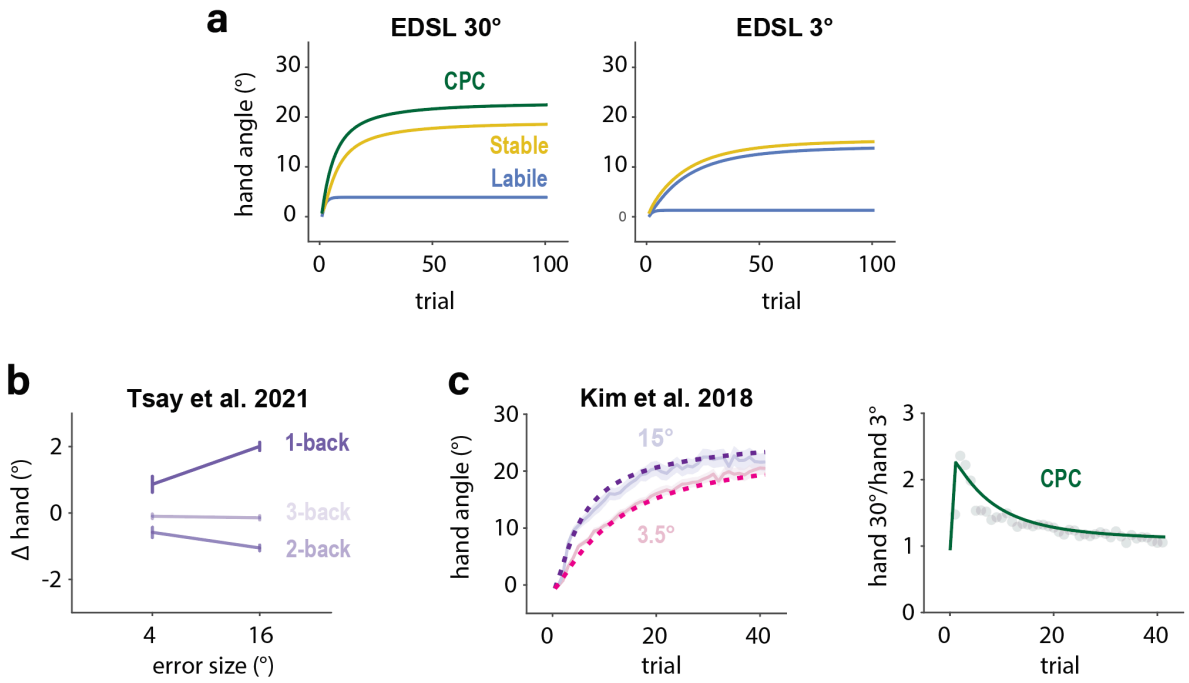


Fig. S6 Learning rates of the labile and stable processes are modulated in a similar way by error size. a) Predicted time course of state of stable and labile processes in Exp 7. Both processes are attenuated in the 3° condition compared to the 30° condition. **b)** Effect of error size on labile processes in an in-person set up, estimated from Exp 3 of Tsay et al[62] in which clamp size was 4° or 16°. **c)**, Learning functions (left) and the ratio between the two error size conditions (right) from Exp 1 of Kim et al[59] with clamp sizes of 3.5° or 15°. Dotted lines show predictions of the CPC model using the learning rate measured by Tsay et al[62]. Shaded area and error bars indicate standard error.

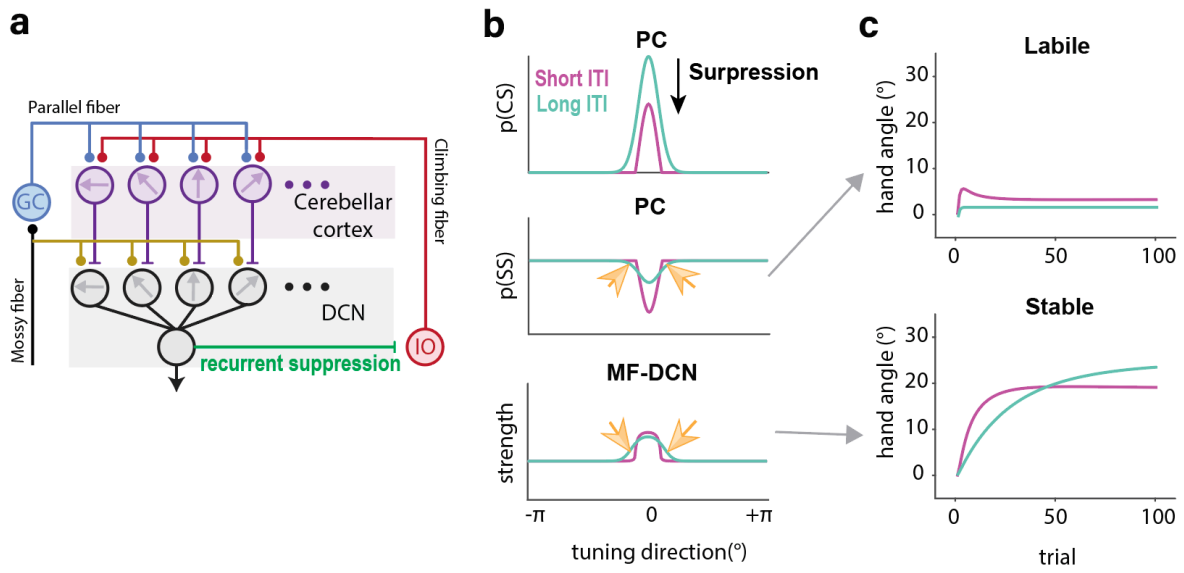


Fig. S7 Revised CPC with recurrent inhibitory pathway. The original CPC predicts that the asymptote should be lower in the long ITI condition compared to the short ITI condition because the latter includes a labile component. However, as shown in Fig 7, the asymptote is similar in the two ITI conditions. This observation motivated a revision to the CPC model in which the DCN sends a recurrent inhibitory signal to the inferior olive. **a)** Model schematic. DCN-IO inhibition suppresses the error signal to the DCN and cerebellar cortex. This suppression is generic given that the output of the DCN integrates activation across directionally tuned units. **b)** When the inter-trial-interval is short, the CS response is suppressed (top). Note that the suppression is implemented by subtracting a common value to the IO and thus alters the activation in PCs. On the next trial, SS activation is stronger in the long ITI condition since the PF-PC synapse will have recovered during the ITI (middle). However, there are a subset of tuned elements that in which SS activation is weaker in the long ITI condition (yellow arrows). This weaker activation induces adaptation in DCN units tuned to the same direction (bottom). **c)** State of the labile and stable processes over the course of a block design under long and short ITI conditions. The change in the labile process is smaller in the long ITI condition due to forgetting. The stable process is also smaller in the long ITI condition because SS activity at the preferred error direction will dominate learning. However, the long

ITI condition induces adaptation in neurons with sub-preferred error directions, resulting in larger adaptation late in training.

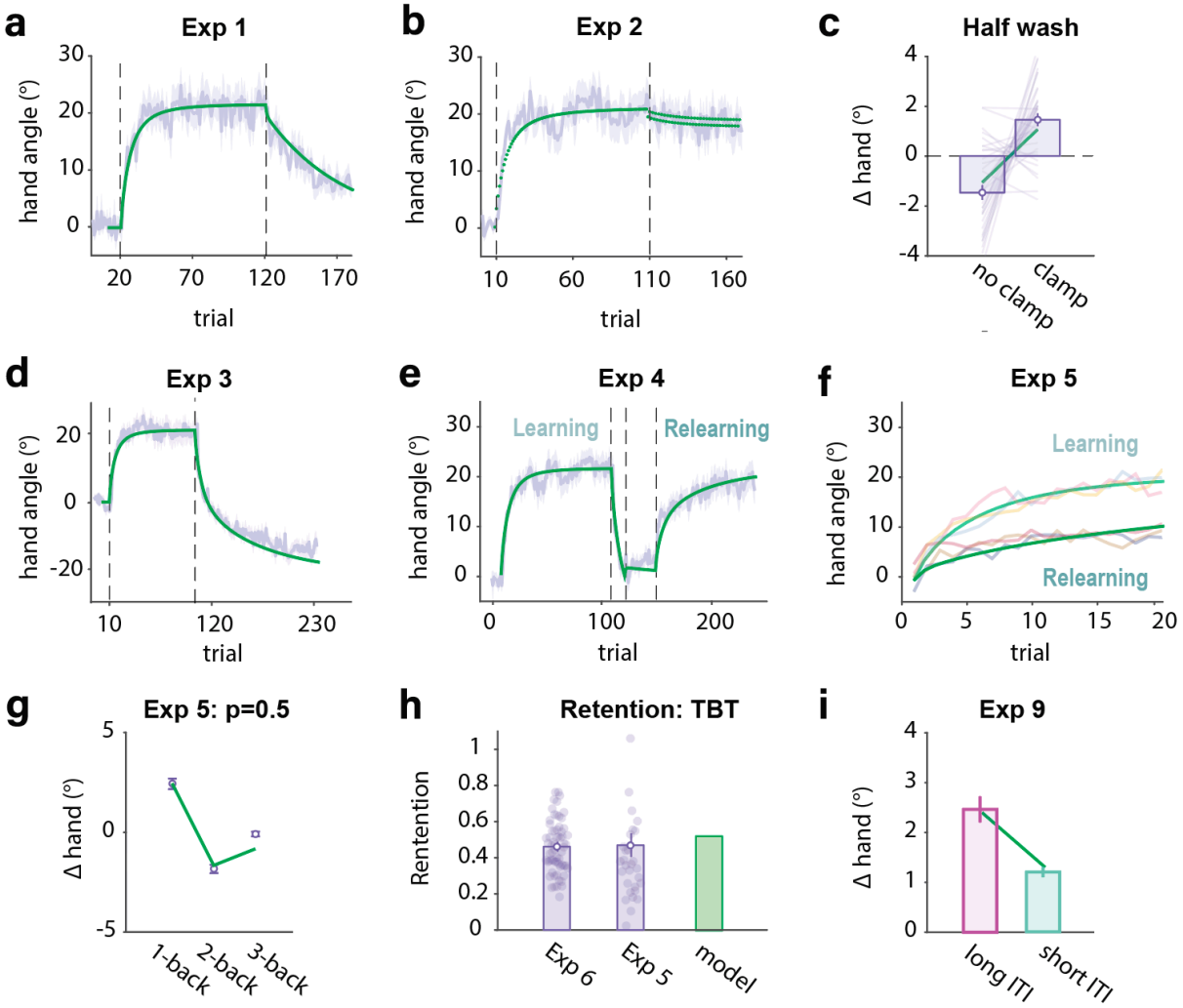


Fig. S8 Revised CPC model provides a good fit for the key results for all of the experiments. Dark green line depicts model prediction. Error bars (c, g, h, i) and shaded areas (a, b, d, e) indicate standard error.

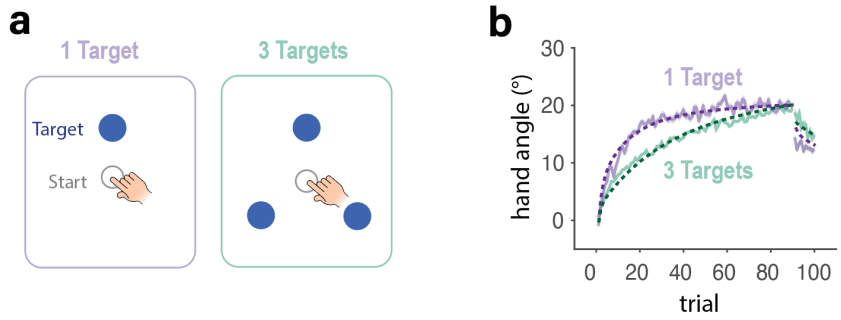


Fig. S9 Revised CPC model accounts for effect of number of target locations on adaptation. a) In Tsay & Irving[89], participants were trained with either one target or three targets. In both conditions, participants reached to a single target during the washout block. **b)** Learning functions for the target location probed during washout. The 3-target condition showed slower learning but a larger aftereffect. Adding more targets is effectively akin to imposing a long ITI since successive reaches to a given target are separated by reaches to the other two locations; thus, there is more forgetting but stronger retention due to reduced contribution of labile process. Shaded area in b indicates standard error. Dash lines indicate the predictions of RSCPC.

1136 Table S1. Comparison of the CPC model and other models of sensorimotor adaptation.
1137

	CPC	Dual State Space[18]	MoE[57]	COIN[29]	Credit Assignment[58]
Minimal attenuation in half washout (Exp 2)	✓	✗	✗	✓	✓
Anterograde interference (Exp 3)	✓	✓	✗	✓	✗
Attenuation in relearning (Exp 4)	✓	✗	✗	✗	✗
Attenuation with opposite errors but being invariant to error consistency (Exp 5)	✓	✗	✗	✗	✗
Different retention rates for trial by trial and block designs	✓	✓	✗	✓	✓
Fast single trial learning to random perturbation (Exp 6)	✓	✓	✗	✗	✓
Fast single trial learning around the asymptote (Exp 2)	✓	✗	✗	✗	✗

1138

1139 Table comparing CPC and other models of sensorimotor adaptations on set of core phenomena (rows). In
1140 evaluating each of the models, we used an implementation based on that presented in the associated
1141 paper (recognizing that a reasonable variant might be possible to capture more of the phenomena). The
1142 listed outcomes are described in the text with the exception of the credit assignment model (Kording et
1143 al., 2017). The credit assignment model assumes that the agent performs Bayesian inference to
1144 decompose the observed error into perturbation sources that vary across different time scales and
1145 estimates the optimal policy to compensate for them.