Cerebro-cerebellar networks facilitate learning through feedback decoupling

Ellen Boven[1,2#], **Joseph Pemberton**[1#], **Paul Chadderton**[2], **Richard Apps**[2], **Rui Ponte Costa**[1*]

[1]Bristol Computational Neuroscience Unit, Intelligent Systems Lab, Faculty of Engineering, University of Bristol, Bristol, BS8 1TH, United Kingdom; [2]School of Physiology, Pharmacology and Neuroscience, Faculty of Life Sciences, University of Bristol, Bristol, BS8 1TH, United Kingdom

Abstract Behavioural feedback is critical for learning in the cerebral cortex. However, such feedback is often not readily available. How the cerebral cortex learns efficiently despite the sparse nature of feedback remains unclear. Inspired by recent deep learning algorithms, we introduce a systems-level computational model of cerebro-cerebellar interactions. In this model a cerebral recurrent network receives feedback predictions from a cerebellar network, thereby decoupling learning in cerebral networks from future feedback. When trained in a simple sensorimotor task the model shows faster learning and reduced dysmetria-like behaviours, in line with the widely observed functional impact of the cerebellum. Next, we demonstrate that these results generalise to more complex motor and cognitive tasks. Finally, the model makes several experimentally testable predictions regarding (1) cerebro-cerebellar task-specific representations over learning, (2) task-specific benefits of cerebellar predictions and (3) the differential impact of cerebellar and inferior olive lesions. Overall, our work offers a theoretical framework of cerebro-cerebellar networks as feedback decoupling machines.

Introduction

Learning ultimately depends on environmental feedback [1,2]. To learn efficiently animals and humans must make good use of this feedback to update their internal models of the world [3,4]. However, external sensory feedback is inherently delayed and incomplete, thereby reducing the rate and extent of learning in neuronal circuits [3]. These observations suggest that the brain may employ a general mechanism to facilitate learning when external feedback is not readily available.

The cerebellum is a region of the brain specialised in building predictive models [4,5]. In the classical view, the cerebellum learns predictive internal models on the motor domain [5–10]. Consistent with this view are a large body of experimental observations for which cerebellar dysfunction causes motor learning deficits. However, more recently, cerebellar dysfunction has also been associated with impaired language processing, cognitive associative learning and working memory [11–15]. Moreover, an increasing body of behavioural [12,14,16–20], anatomical [21,22] and imaging [23] studies alludes to a role of the cerebellum in cognition in animals and humans. Taken together, these studies suggest that the cerebellum learns internal models for both motor and non-motor functions in line with the proposed *universal functional* role of the cerebellum across the brain, including the cerebral cortex [9,24–26].

Despite growing experimental evidence there are no specific computational models aiming to capture the functional roles of cerebro-cerebellar interactions during learning of motor and non-motor tasks. Building on recent deep learning developments we theorise that the cerebellum predicts future cerebral feedback signals given current cerebellar activity. This feedback predicted by the cerebellum is then sent back to the cerebral network to drive learning.

[#]These authors contributed equally (order is alphabetical).

*For correspondence: rui.costa@bristol.ac.uk

Specifically, we model a given cerebral area as a recurrent neural network [27–30] which receives feedback predictions from a feedforward, cerebellar, network [6,7]. This view of cerebro-cerebellar interactions is in line with the classical forward models of cerebellar function [6,7], in that in our model the cerebellum makes forward predictions (i.e. generates cerebral feedback predictions) given current cerebral activity.

We test our model on a range of sensorimotor, pattern recognition and visual-language tasks. Using these tasks we demonstrate that cerebellar predictions conveyed to the cerebral cortex facilitate learning. Moreover, models without a cerebellar component exhibit slower learning and dysmetria-like behaviours, consistent with a wide range of behavioural observations [11,14,31,32]. Our results indicate that the cerebellar-mediated facilitation of cerebral learning relies on the ability of the cerebellum to provide effective cerebral feedback predictions. Finally, we make several experimentally testable predictions regarding cerebro-cerebellar representations, task-specific temporal feedback, cerebro-cerebellar activity coupling and the different contributions of cerebellar output and inferior olive lesions for task learning.

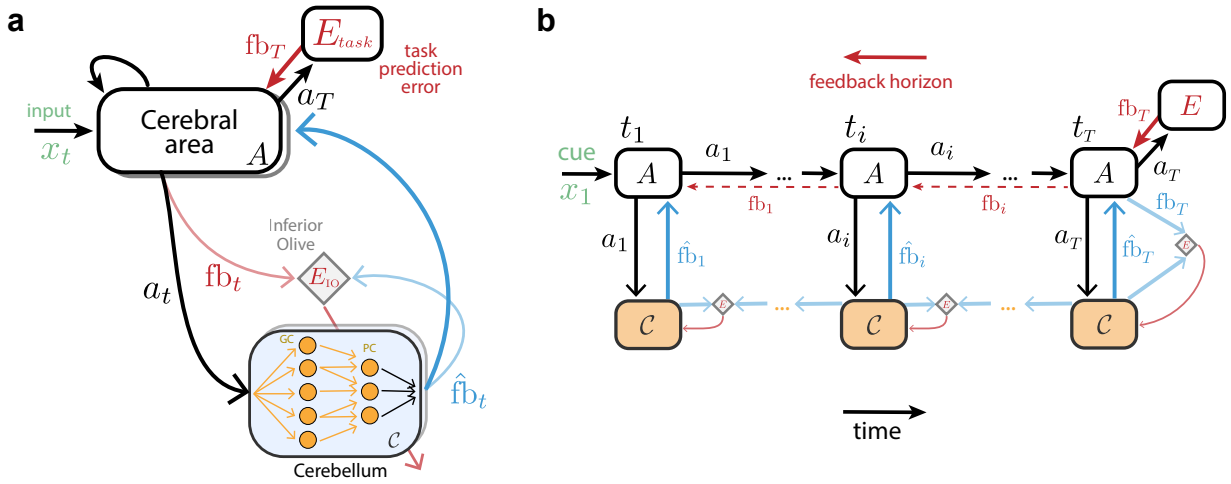


Figure 1. Cerebro-cerebellar networks as feedback prediction machines. (**a**) A recurrent cerebral cortical network *A* learns through feedback given by a task-specific prediction error module E_{Task} computed at the end of a task fb_T (top red arrow). The cerebellum aims to continuously predict the feedback expected by the cerebral network $\hat{\text{fb}}_t$ (blue) given current cerebral activity a_t (black). The cerebellar network (i.e. granule cells; GC and Purkinje cells; PC) learns through prediction errors (bottom red arrow) computed at the inferior olive (diamond) by comparing predicted feedback $\hat{\text{fb}}_t$ with actual feedback fb_t (light blue). Shaded boxes represent multiple cerebral areas and cerebellar modules that may be interacting in parallel (see Fig. S1 for the same framework applied to decoupling across multiple brain areas). (**b**) Example of cerebro-cerebellar model unfolded in time in which the cerebral network learns to associate a cue given at t_1 (x_1, green) with feedback received at the end of the task, t_T (cf. Fig. 2). At the end of the task the cerebral network *A* receives external sensory feedback fb_T (red), which is transmitted to the cerebellar network as cerebral feedback fb_T (light blue). Here we highlight a case of cerebral feedback horizon stopping at the end of the task *T*, but feedback may also be available earlier in the task (dashed red arrows). The cerebellum generates cerebral feedback predictions $\hat{\text{fb}}_T$ (blue) given cerebral activity a_T (black), and learns using inferior olive (diamond) error signals (red arrow). Before t_T cerebral feedback may not be readily available, thus the cerebellum learns through self-predictions. In this case the inferior olive (diamond) compares old cerebellar predictions (e.g. $\hat{\text{fb}}_i$) with the new one (e.g. $\hat{\text{fb}}_T$) to generate cerebellar learning signals (red arrow; see main text and Methods for details).

Results

A systems-level computational model of cerebro-cerebellar interactions

In order to understand how cerebellar computations may shape cerebral processing, we introduce a cerebro-cerebellar systems-level model based on a recent deep learning algorithm [33]. In line with previous work we model a given cerebral cortical area *A* as a recurrent neural network (RNN) [27–30] which is coupled with a cerebellar module *C* – cerebro-cerebellar RNN (ccRNN). We model the cerebellar module as a simple feedforward network *C* (Fig. 1a) in line with the cerebellar architecture [6,7,9]. The input layer of the cerebellar network receives cerebral cortical activity *a* and models the Granule cells (GCs), which project to the output layer, modelling the Purkinje cells (PCs) that provide cerebellar predictions back to the cerebral cortex (Methods). To capture the dimensionality expansion observed between cere-

bral and cerebellar networks[34,35] we constrain our model with $M \gg N$, where M corresponds to the number of GCs and N the number of cerebral neurons and use the same ratio found experimentally $\frac{M}{N} \sim 4$.

We study the behaviour of our model in a range of tasks. To train the model we use a prediction error function E_{task} which compares the model output with task-specific external feedback. Using standard gradient descent methods we generate feedback signals of a specific temporal horizon (see example of a RNN unrolled in time in Fig. 1b), fb_t, which is then used to update the RNN input and recurrent weights (Fig. 1a; see Methods). For computational efficiency and in line with previous models we use a time-discrete approximation of time-continuous RNN models[28].

Following our theoretical proposal the cerebellar module C continuously learns to predict cerebral feedback fb_t given cerebral cortical activity a_t. The cerebellar network is optimised through error signals computed by comparing the actual cerebral feedback fb_t at time t with the cerebellar predicted feedback $\hat{\text{fb}}_t$. We postulate that this comparison is done in an inferior olive-like structure, $E_t^C = (\text{fb}_t - \hat{\text{fb}}_t)^2$, that generates error signals which are used to optimise the cerebellar network (see Methods). However, similar to the external feedback, actual cerebral feedback is not always available, which would impact the ability of the cerebellar network to learn online to produce effective feedback signals. To circumvent this problem we propose that the cerebellum learns using its own feedback predictions when cerebral feedback is not available (Fig. 1b)[33]. This leads to the following target feedback $\bar{\text{fb}}_t \sim \text{fb}_t + C(a_{t+1})$ where fb_t is the true cerebral feedback and $C(a_{t+1}) = \hat{\text{fb}}_{t+1}$ is a self-prediction term which enables the cerebellum to learn online (see full details in Methods).

Cerebro-cerebellar model facilitates learning in a simple sensorimotor task

Inspired by classical sensorimotor studies in the cerebellum, we first test a simple visuo-motor task[11,31,32,36,37]. In this task the model must draw a straight line in a two-dimensional space towards one of seven target locations given a target-specific cue at the start of the task (Fig. 2a, top left). We train a cerebro-cerebellar RNN (ccRNN) and a cerebral-only RNN (cRNN) to perform this task (see full details in Supplementary). To train the models we provide teaching feedback by comparing the cerebral network output with the optimal trajectory (i.e. a straight line between starting and end points; Fig. 2a). In addition, this feedback is delayed with respect to the initial cue and incomplete (i.e. only available every few time steps). This models a more realistic setting in which task feedback is not always readily available. When this feedback is available at time t we calculate the prediction error as $E_{\text{task}} = (l_t - \hat{l}_t)^2$, where l_t and $\hat{l}_t$ denote the desired and current model two-dimensional trajectory (i.e. set of feedback points; cf. Fig. 2 schematic), given by a linear readout on the network activity a_t (Methods). In particular, here we consider a feedback interval at every other time step for both cRNN and ccRNN (but see Fig. 4 for more general cases).

During learning the ccRNN model achieves near-zero error after a relatively small number of training sessions, while the cRNN, which lacks the cerebellar component, also learns but more slowly and with higher variability (Fig. 2b). These observations are in line with a large body of cerebellar experiments[11,31,32]. In addition, we also observe differences at the level of model output trajectories. While the ccRNN produces smooth and straight trajectories, the cRNN displays a much more variable trajectory towards all targets (Fig. 2b). Due to the sparse task feedback in the absence of a cerebellar network, the cRNN is not able to learn a correct trajectory in points for which there is no direct feedback thus overshooting the target trajectory. In cerebellar patients, this effect is referred to as dysmetria[38] which in the motor domain results in ataxia. Ataxia is the lack of coordination and fine control during voluntary movements, a defining symptom resulting from cerebellar malfunction[11,38]. To evaluate the degree of dysmetria-like output in our models we measure the error between the model output and the optimal trajectory (i.e. a straight line in this case; see Methods). When applying this measure, the ccRNN shows a clear reduction in ataxia-like behaviour compared to cRNN (Fig. 2c).

To highlight the conditions for which the cerebellum may facilitate learning in cerebral networks we test different lengths of cerebral feedback horizon (Methods). Our results show that the ccRNN only facilitates learning for short to medium feedback horizons (<50%, Figs. 2d, S2). These results suggest that the cerebellum is particularly important for cerebral learning in conditions in which cortical networks do not have internal effective feedback available for learning. This is consistent with experimental observations showing that the cerebellum becomes more important in the presence of challenging task conditions for which cerebral feedback might be short[39]. In contrast, for long cerebral feedback, having a cerebellar module harms learning. In this case the cerebral network has the level of feedback required to learn effectively, thus the noise inherent in the cerebellar feedback can impair learning. This observation suggests that the brain may use intermediate brain structures, such as the thalamus and the pons to

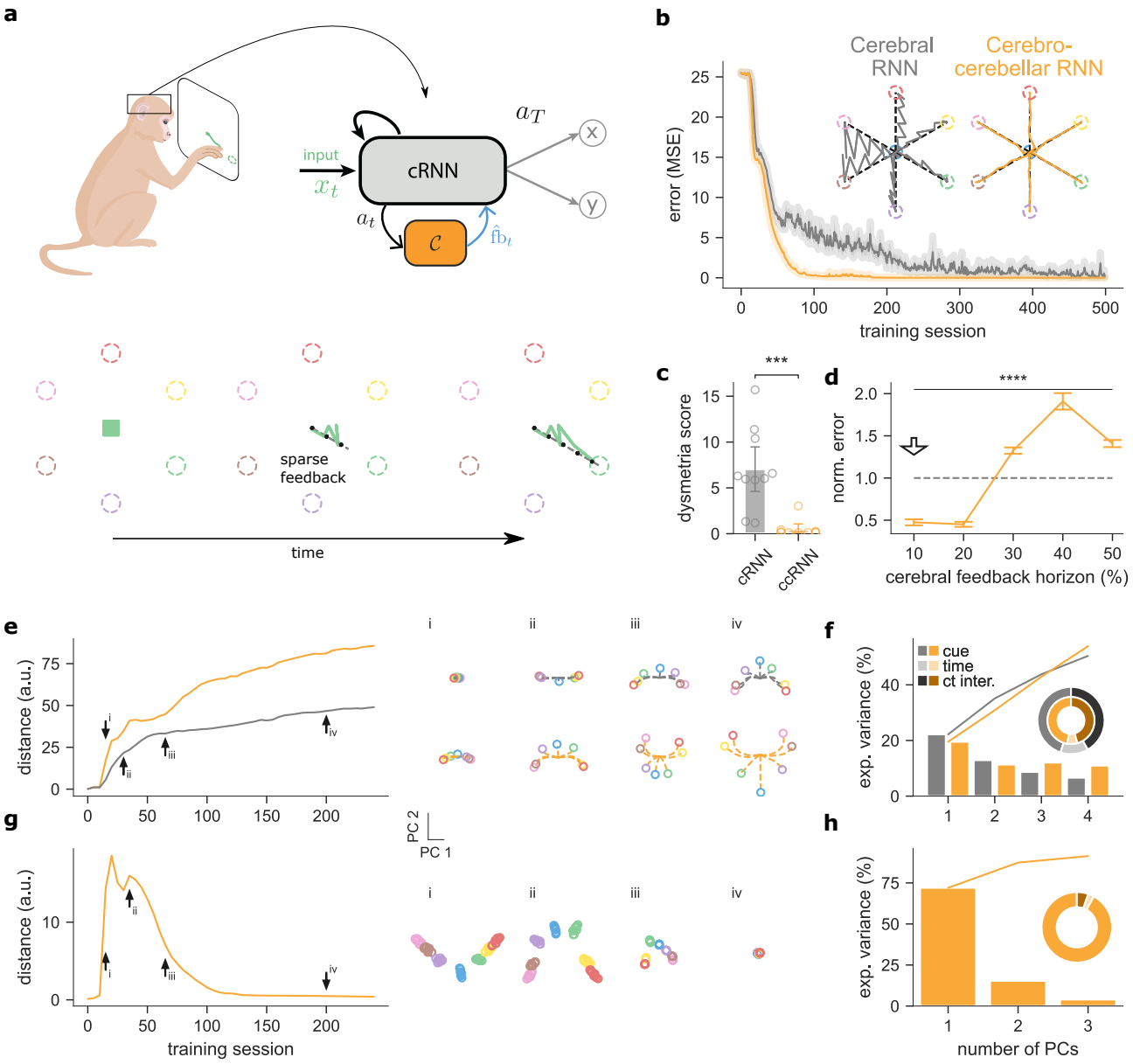


Figure 2. Cerebro-cerebellar model improves learning in a simple line drawing sensorimotor task. (**a**) Schematic of a macaque monkey performing a simple line drawing task (top left). A cerebro-cerebellar RNN (ccRNN) in the macaques brain receives the cue-specific input and learns to produce the desired trajectory (top right). There are 6 possible targets (coloured dashed circles) and feedback (dashed black line) is provided at a regular interval (bottom; see Methods). In the example shown the model must draw a straight line towards the green target. (**b**) Error between model output and desired target trajectories for cerebellar RNN (gray, cRNN) and cerebro-cerebellar RNN (orange, ccRNN). Insets: Model trajectory produced for all cues after learning. (**c**) Dysmetria score for cRNN and ccRNN. The dysmetria score quantifies how smooth the movement is after learning (Methods). (**d**) Normalized model mean squared error (MSE) after learning for different cerebral feedback horizons. Feedback horizon is denoted as percentage of the total task sequence. Arrow indicates feedback horizon used by the cerebral network in the other panels. (**e**) Euclidean distance between the two leading cue principal components for the recurrent neural network in both the cRNN (grey) and ccRNN (orange) models. Arrows highlight training sessions of cue-specific principal components (PCs) plotted on the right for early (i), early-mid (ii), mid (iii) and late (iv) learning, for both cRNN (top) and ccRNN (bottom). (**f**) Explained variance of the RNN for both models cRNN (gray) and ccRNN (orange). Bar plot shows explained variance for the top five cue-specific PCs. Circular plot shows the total explained for cue (medium-dark colours), time (light colours) and cue-time interaction (dark colours) task variables. (**g**) Euclidean distance between the different cue-specific two-dimensional components for the cerebellar network (orange, ccRNN model). Arrows indicate training sessions highlighted on the right as in (e). (**h**) Explained variance of the cerebellar network as in (f). ***: p<0.001, ****: p<0.0001. Error bars represent mean ± SEM across 10 different initial conditions of the model.

gate cerebro-cerebellar interactions depending on task properties (see Discussion).

Next, to gain insight into how cerebral and cerebellar neuronal representations evolve jointly during learning, we use a dimensionality reduction method (demixed principal component analysis (PCA); see Methods). Demixed PCA (dPCA) enables us to extract low-dimensional neuronal representations that capture maximum variance across task variables. First, we focus on the two most informative cue-specific principal components using the neural activities of the recurrent neural network for both cRNN and ccRNN (see all components in Figs. S3,S4 and S5). Next, we calculated the two-dimensional Euclidean distance across the 7 different possible cues (Methods). Our results show that the ccRNN cerebral network is characterised by a stronger increase in separation of stimulus components over learning when compared to the cRNN cerebral network (Fig. 2e). To contrast task-specific components with general temporal information, we compare the level of cue-specific and time-specific explained variance in both models. ccRNN captures overall more cue-specific explained variance when compared with cRNN (Fig. 2f) which demonstrates that ccRNN encodes more task-relevant information, which requires the model to associate the cue information with specific output trajectories. Next, we applied dPCA to the activity of cerebellar neurons. Since the cerebellar module facilitates cue-to-target learning we expected cerebellar representations to be mostly dominated by task-specific information. This is indeed what we find, our results show that the distance between cue-related components is stronger during periods of high learning (Fig. 2g; compare with Fig. 2b), and that most of the variance is explained by cue-specific PCs (95.4%; Fig. 2h).

Overall, our results suggest that in the context of a simple sensorimotor task, cerebellar-mediated decoupling of cerebral feedback enables faster learning and smoother motor trajectories. In addition, it makes a number of experimentally testable predictions about the evolution of task-specific cerebro-cerebellar representations throughout learning.

Cerebro-cerebellar model improves learning in complex sensorimotor and discrimination tasks

Next, to test whether the results from the simple visuomotor task generalise to more realistic settings we explore a range of more advanced sensorimotor tasks. We introduce two tasks in which the models are trained to draw digits given complex spatiotemporal sensory inputs. For these tasks we build on a standard machine learning dataset consisting of 10 (from 0 to 9) two-dimensional handwritten digits (see example in Fig. 3a; MNIST dataset [40]). In contrast to the previous task in which sensory input was only provided at the start of the task, here the model receives a part of a handwritten digit at any given point in time (i.e. a row of 28 pixels; see Methods). We refer to this task setting in which input is provided over time as *online*. Given this input then we consider two task variants (Fig. 3a) in which the model has to either (i) draw a straight line (online line drawing (LD) visuomotor task) or (ii) draw a digit (online digit drawing (DD) visuomotor task. Both tasks provide a more realistic model of drawing tasks (Fig. 2) in which lines must be drawn given complex continuous sensory input. As in the previous task we consider cases of sparse task feedback.

As in the simple visuomotor task, here the ccRNN learns faster (Fig. 3b) than cRNN while showing a strong reduction in dysmetria-like trajectories (Fig. 3c). The ccRNN also facilitates learning when in the presence of short to medium feedback horizon in the cerebral network (Fig. 2d), and we find that dysmetria-like trajectories are reduced in the ccRNN model (Fig. 3c).

To test whether our observations in the sensorimotor tasks generalise to other task domains we train the model in a visual discrimination task. In this task the model receives the same handwritten digits presented sequentially over time but now must discriminate between the 10 classes of digits (online visual discrimination task, Fig. 3a). In line with the results in the visuomotor tasks, we find that ccRNN also facilitates learning in this task, achieving higher accuracy after only 10 training sessions (Fig. 3b). Here we use the certainty the model has about the current class as a measure of dysmetria of thought [41] (see Methods). Similarly to the tasks above, we find that dysmetria-like behaviours are reduced in the ccRNN model, which in this case shows that model produces more accurate decisions (Fig. 3c). Finally, in line with previous tasks a cerebellar module facilitates learning in the presence of weak cerebral feedback (Figs. 3d, S6). These results are in line with the growing number of studies implicating the cerebellum in sensory discrimination and decision making tasks [19,42,43].

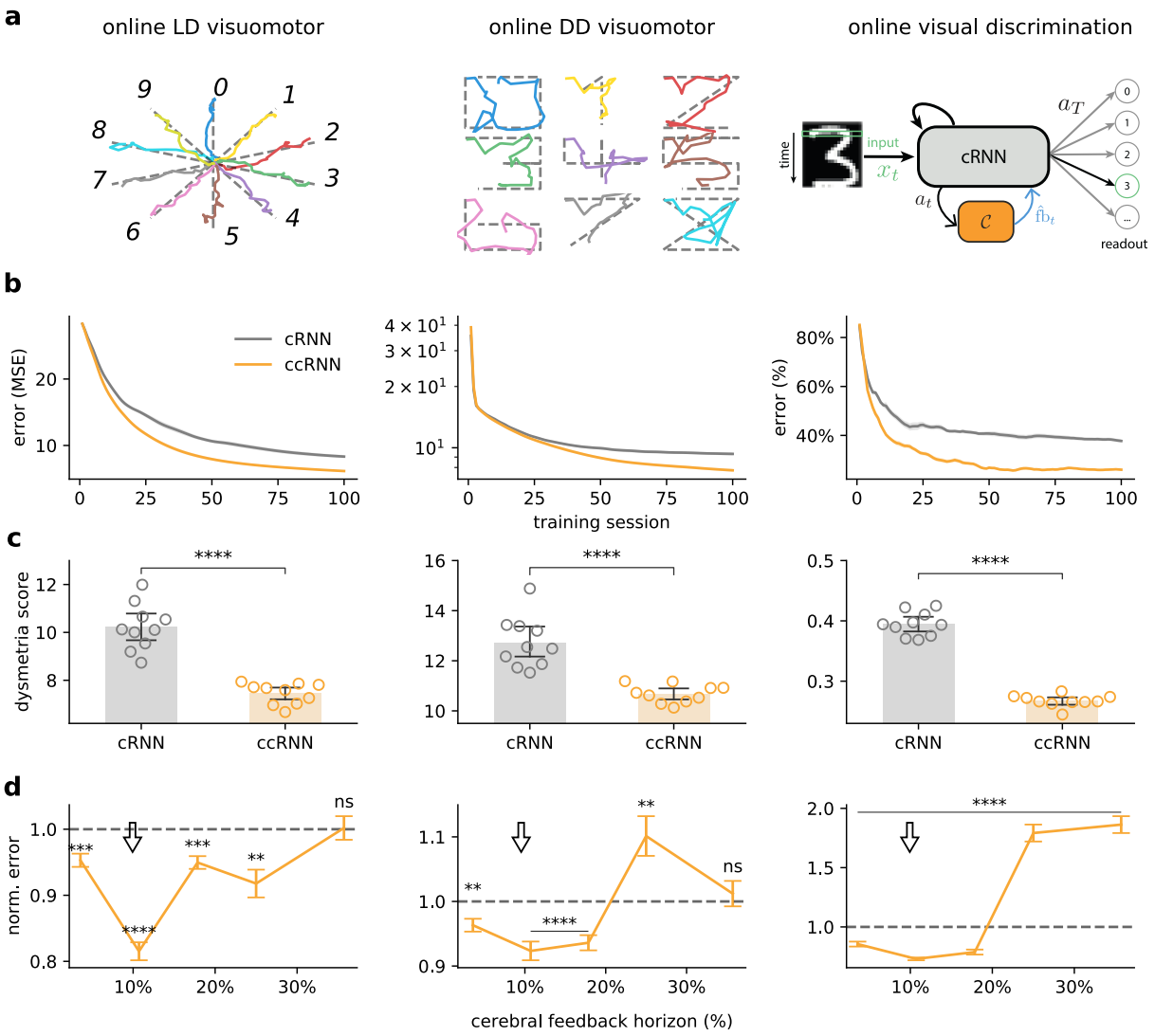


Figure 3. Cerebro-cerebellar model improves learning in online complex sensorimotor and sensory discrimination tasks. (**a**) Model behaviour across three tasks using a dataset of handwritten digits, each presented sequentially to the network (Methods and main text). Online line drawing (LD) visuomotor task: given temporally varying visual input the model is trained to draw a straight line (top left). Online digit drawing (DD) visuomotor task: given temporally varying visual input the model is trained to draw a digit following a template (top middle); Target trajectories are in dotted grey and model output is coloured by digit. Online visual discrimination task: pattern recognition variant in which the model is trained to discriminate between 10 different digits given as input sequentially (a row at a time; green box; top right). (**b**) Learning curves for the three tasks for both cerebral RNN (gray, cRNN), cerebro-cerebellar RNN (orange, ccRNN). The cerebral network in all tasks uses approx. 10% of the cerebral feedback horizon (cf. d). (**c**) The dysmetria score quantifies the irregularity in movement during the testing phase of the model (online LD and DD visuomotor tasks) or the uncertainty in the sensory discrimination (online visual discrimination task). (**d**) CcRNN model performance relative to cRNN across different degrees of cerebral feedback horizon. ns denotes no significance (p=0.921 in the online LD visuomotor and p=0.567 in the online DD visuomotor). Arrow indicates the feedback horizon used in (b). **: p<0.001 ***: p<0.0001, ****: p<0.0001. Error bars represent mean ± SEM across 10 different initial conditions of the model.

Cerebellar-mediated learning facilitation depends on task feedback interval

In sensorimotor tasks there are physiological constraints inherent to animals and humans which impose limits on the rate at which external feedback is available[44–46]. To determine the rate of external feedback for which cerebellar predictions are most valuable we trained the model in two tasks (simple LD and LD visuomotor tasks) with a range of external feedback intervals. This feedback interval defines the rate at which external feedback is available for learning, resembling sensorimotor feedback which is typically sporadic rather than continuous[11,47,48]. We find

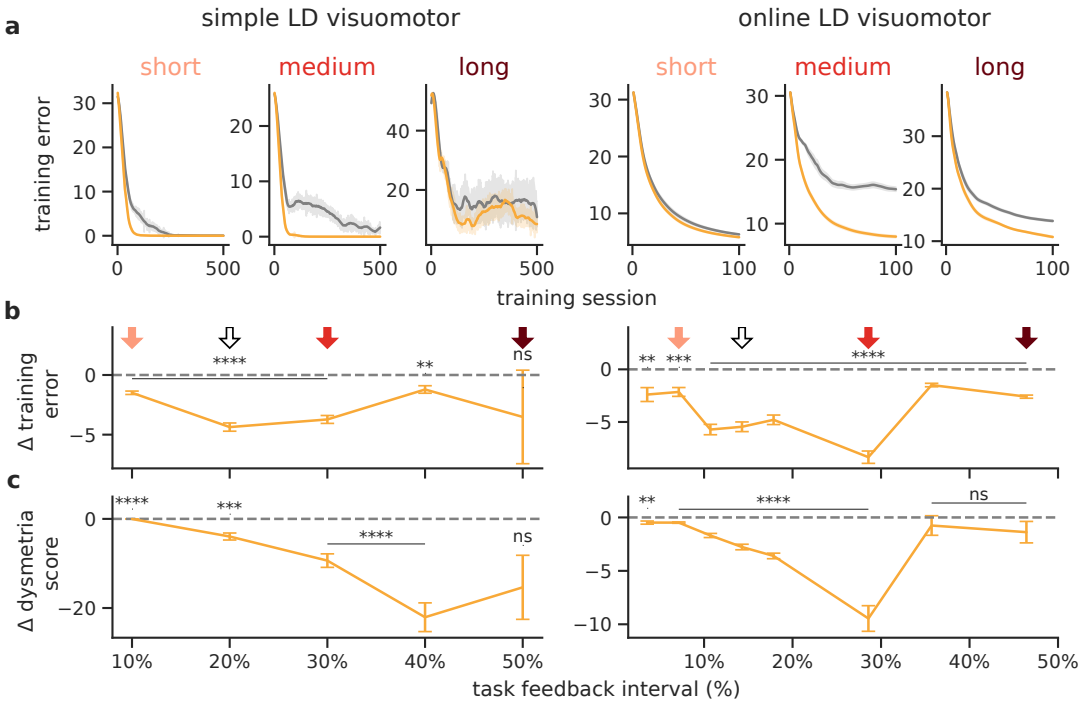


Figure 4. Cerebellar-mediated facilitation of learning depends on task feedback interval. (**a**) Learning curves for short (light red), medium (red) and long (dark red) levels of feedback interval for both the simple and online LD visuomotor tasks and both models cRNN (gray) and ccRNN (orange). Degrees of redness (**b**) Difference in task error between ccRNN and cRNN for varying degrees of task feedback intervals (not significance, p=0.406). Degrees of red in arrows indicate the respective interval in (**a**) while the white arrow indicates the feedback interval used in Fig. 2 and Fig. 3, respectively. Task feedback interval given as a percentage of the total task time. (**c**) Difference in dysmetria score between for varying degrees of task feedback interval (not significance, p=0.0577 for simple LD and p=0.444 (40%), p=0.209 (50%) for online LD). **: p<0.01, ***: p<0.001, ****: p<0.0001. Error bars represent mean ± SEM across 10 different initial conditions.

that when external feedback is given at short intervals there is little advantage of the feedback predictions from the cerebellar component for both the simple LD and online LD visuomotor tasks (Fig. 4a,b). When the interval between external sensory feedback is increased, the benefits of the cerebellar-to-cerebral feedback predictions for learning in the ccRNN model become clear. In contrast, for long feedback intervals the feedback is too infrequent for both cRNN and ccRNN to be able to successfully learn the task. Next we evaluate the degree of dysmetria using the metrics introduced above. We observe qualitatively similar results: a model without a cerebellar network (cRNN) exhibits more variable trajectories for medium to long task feedback intervals (Fig. 4a). These results imply that whether cerebellar-to-cerebral feedback is beneficial for learning and leads to dysmetria-like behaviours depends on the rate of task feedback.

Similarity between cerebellar and cerebral feedback is task and learning dependent

The cerebro-cerebellar facilitation of learning shown above depends on the ability of the cerebellum to provide the cerebral network with effective feedback predictions. To study the level of similarity between the cerebellar predicted feedback and the theoretically optimal cerebral feedback as provided by gradient descent methods, we calculated the cosine similarity between cerebellar predictions and the optimal cerebral feedback in a range of tasks (Methods).

First, we measure the cosine similarity for tasks in which external sensory feedback is only provided at the end of the task – a variant of the simple LD task with feedback only at the end and the online visual discrimination. This task setup allows for an easier interpretation of the similarity between cerebellar and cerebral feedback which should decay gradually from the end to the beginning of the task sequence. Indeed, we observe that the cerebellar-cerebral feedback similarity is higher closer to the point in which external sensory feedback is available (i.e. end of the task; Fig. 5a,b top; cf. Figs. 2, 3) and remains high over learning in particular for later points in the task (Fig. 5a,b bottom).

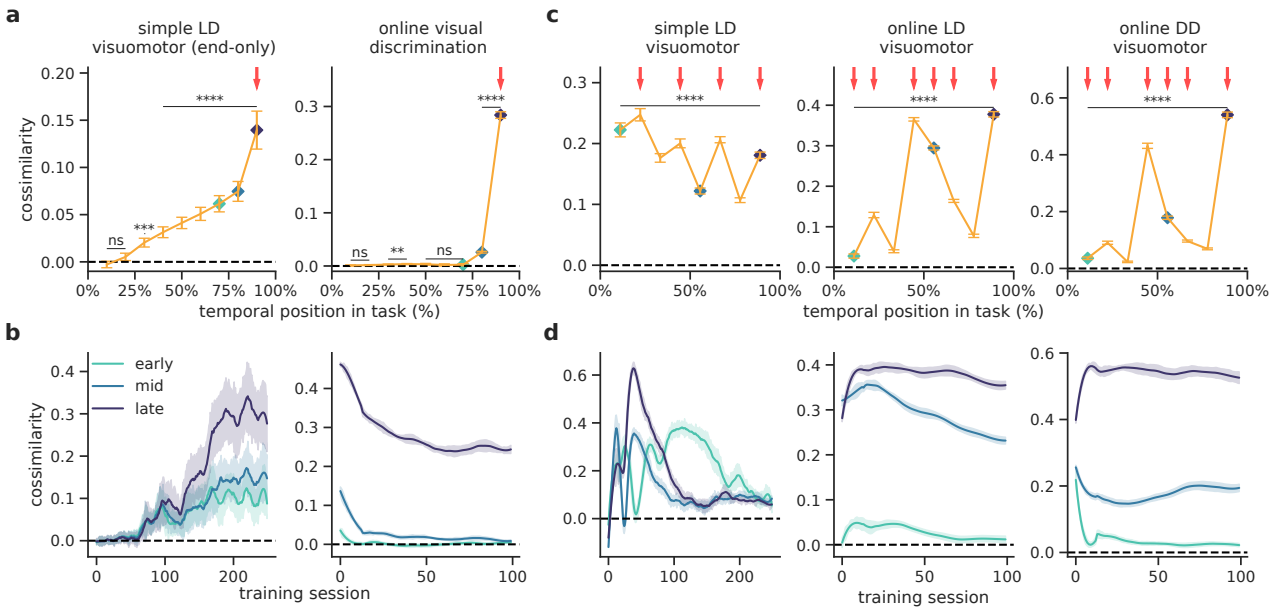


Figure 5. Similarity between cerebellar and cerebral feedback is task and learning dependent. (**a**) Cerebro-cerebellar cosine similarity throughout tasks sequences which do not require intermediate external feedback: simple line drawing with feedback only at the end of the task (LD end-only) and online visual discrimination (n.s. simple LD visuomotor p=0.212 (0%), p=0.520(25%), n.s. online LD visuomotor p=0.312 (0%), p=0.06 (25%), p=0.067(50%), p=0.386(60%). Here and in subsequent panels red arrows indicate points in which external feedback is available. Cosine similarity throughout the tasks is calculated across all training sessions (see Methods).(**b**) Cerebro-cerebellar cosine similarity over learning for three time points in the task: early (turquoise), mid (blue) and late (purple) in the task (cf. (a)). (**c**) Cerebro-cerebellar cosine similarity throughout the sequence for tasks with intermediate external feedback: simple line drawing (LD), online LD, online digitdrawing (DD). (**d**) Cerebro-cerebellar cosine similarity over learning for three different time points in the task (early, mid and late as in (b)). Dashed black line represents zero similarity. **: p<0.01, ***: p<0.001, ****: p<0.0001. Error bars represent mean ± SEM across 10 different initial conditions.

Next, we analyse the cosine similarity for conditions in which external feedback is available throughout the task. For this we consider the same visuomotor tasks as above (simple LD visuomotor, online LD visuomotor and online LD visuomotor). In these tasks we observe more complex dependencies of the cerebro-cerebellar feedback similarity on task properties (Fig. 5c,d). For the simple LD task we observe that the predictions made during earlier points in the task are more similar than those at later points (Fig. 5c). These results suggest that the model is first learning to align later points in the task and gradually learns to adjust earlier points which are closer to the cue-specific information that defines the trajectory that the model must take. Interestingly, this behaviour is less prominent in the two other tasks, online LD and DD visuomotor tasks, that are characterised by relatively more complex task-specific sensory input occurring throughout the task. For these two more complex tasks and in contrast to the simple LD the similarity remains high throughout learning for later time points (Fig. 5d), which reflects the more challenging nature of these tasks and the need to continuously predict feedback as the task is never fully learnt.

These results make non-trivial predictions on when the cerebellum is able to better align with the cerebral feedback, which depend on task complexity, the properties of the task feedback, the exact task position and the learning stage.

Learning shapes cerebro-cerebellar activity coupling

The cosine similarity results show that the cerebellar module learns to predict cerebral feedback. Because the cerebellum maps cerebral activity onto (predicted) cerebral feedback, this suggests changes in the coupling between cerebellar and cerebral neuronal representations throughout learning. To study the degree of cerebro-cerebellar coupling we calculate the pairwise correlations between neurons in the cerebral recurrent neural network and the neurons of the cerebellar network (Methods). Although we observe a relatively small rise in the average cerebro-cerebellar coupling during the first few training sessions, as training progresses, there is a consistent decrease of the correlations (Fig. 6a).

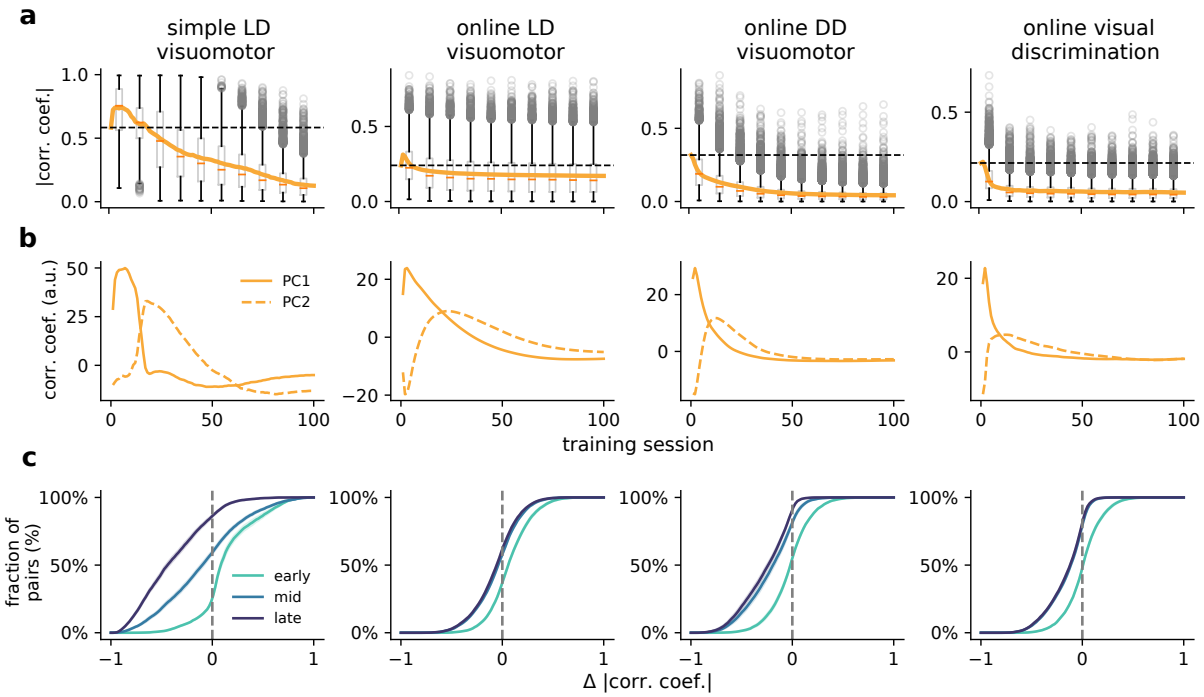


Figure 6. Cerebro-cerebellar neuronal activity coupling over learning. (**a**) Box plot showing the mean and distribution of pair-wise cerebro-cerebellar absolute correlation coefficients over learning for four tasks: simple LD, online LD, online DD and online visual discrimination. Fully fixed ccRNN (i.e. without any form of plasticity in both networks) is given for reference (dashed line). (**b**) Change in first two principal components of cerebro-cerebellar pair-wise correlation coefficients over learning (all components available in Fig. S7). (**c**) Cumulative plot of cerebro-cerebellar pairs with positive and negative changes in absolute correlation coefficients in early (session 1), mid (session 25) and late (session 80) learning. Error bars represent mean ± SEM across 10 different initial conditions.

To study more subtle changes in the correlation structure we use standard principal component analysis of the pairwise correlations (Fig. 6b). The first principal component reflects the changes in the average cerebro-cerebellar coupling (Fig. 6b). The second principal component shows a delayed increase with respect to the first, followed by a sustained decrease in the cerebro-cerebellar coupling (see Fig. S7 for remaining components). These results are consistent with the need for the cerebellum to provide more effective feedback and thus be more coupled in the earlier learning phases. To study learning periods of consistent increases or decreases in coupling as training progresses we tracked the changes in correlations of cerebro-cerebellar pairs in early, mid and late learning (Figs. 6c). We observe that early in learning – when most learning occurs – a large part of the population shows a consistent increase in correlations, but this rapidly changes as learning progresses with only a very small number of pairs showing increases in correlations later in learning.

To better assess the contribution of a plastic cerebellum to the cerebro-cerebellar coupling, we analysed a ccRNN in which the cerebellum does not learn. In this case we can still observe changes in cerebro-cerebellar coupling over learning for some tasks, which reflect changes in the RNN itself, but these are weaker when compared to the normal ccRNN (Fig. S8a). In this case cerebro-cerebellar correlations remain high throughout learning compared to a ccRNN with a plastic cerebellum. This is supported by their low-dimensional representations: whereas a plastic cerebellum leads to principal components that approach near-zero values after the initial learning phase (Fig. 6b, S7), in the case of the fixed cerebellum the principal components continue to fluctuate throughout learning (Fig. S8).

Although our model suggests a long-term decrease decrease in the cerebro-cerebellar activity coupling, it highlights sub-populations which increase their coupling during specific periods of learning. This observation follows from our proposal in that the cerebellum is trained to map cerebral neuronal activity on cerebral feedback which depend on learning.

Differential impact of cerebellar output and inferior olive on learning

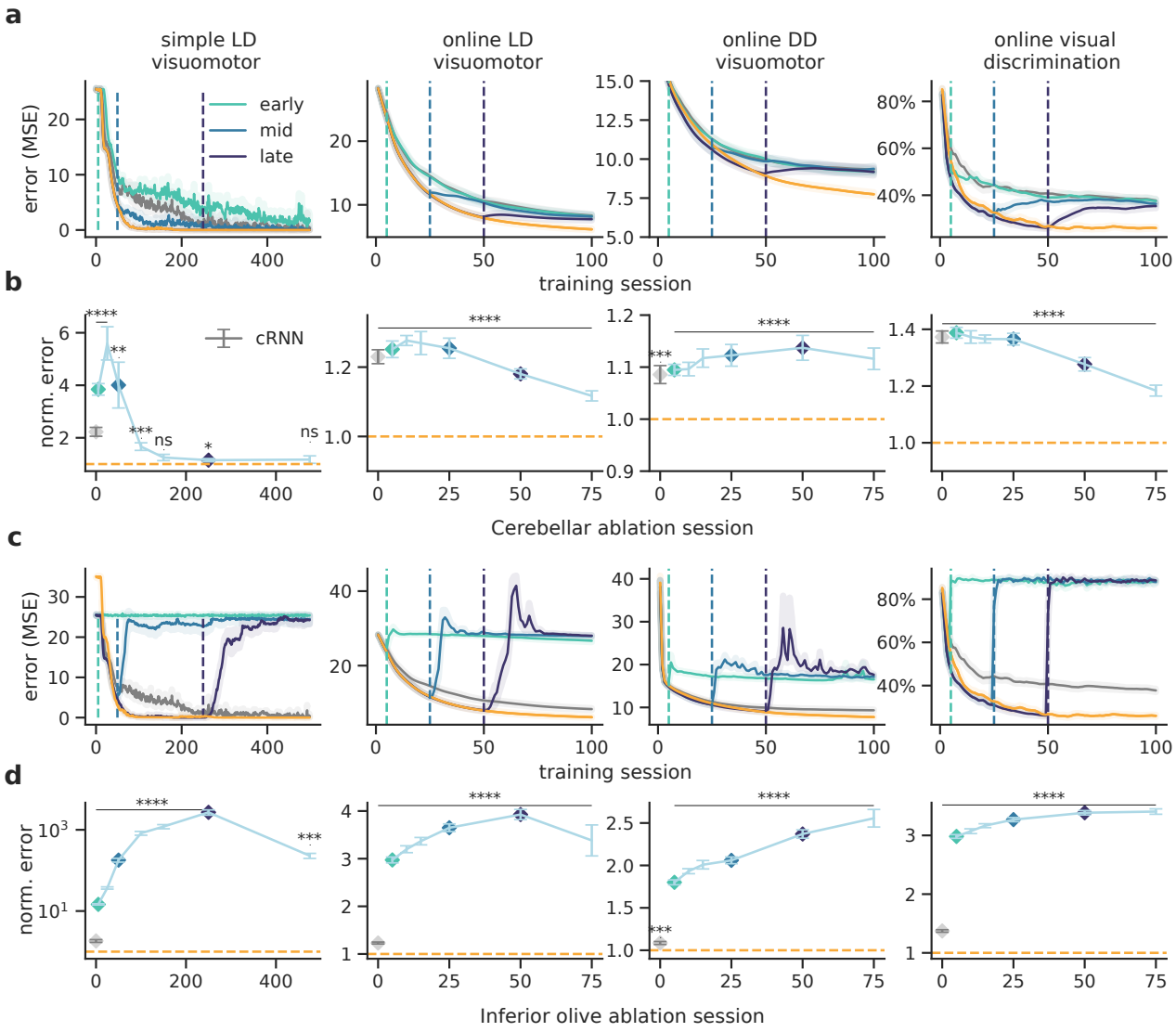


Figure 7. Inactivating cerebellar output and inferior olive have a differential impact on learning. (**a**) Complete cerebellar lesion at different points during learning. Vertical lines represent at which point during training the cerebellar was inactivated in the ccRNN model. In gray and orange show the baseline performances of the cerebral RNN and ccRNN, respectively. (**b**) Normalised error after cerebellar lesion throughout learning with respect to ccRNN (n.s. simple LD visuomotor p=0.062 (session 150), p=0.162 (session 475)). Gray denotes normalised error for cRNN. (**c**) Complete inferior-olive lesion at different points during learning. Vertical lines represent point of lesion of the ccRNN model. In gray and orange are shown the baseline performances of the cerebral RNN and ccRNN, respectively. (**d**) Normalised error after inferior-olive lesion throughout learning with respect to ccRNN. Gray denotes normalised error for cRNN. *: p<0.05, **: p<0.01, ***: p<0.001, ****: p<0.0001. Error bars represent mean ± SEM across 10 different initial conditions.

In experimental neuroscience a common paradigm is to inactivate the cerebellum in order to study its role in learning and behaviour. Here we perform *in silico* lesion experiments to reveal the impact of the modelled cerebellar feedback predictions during learning. First, we test cerebellar output lesions at different points in learning. In all tasks we observe that inactivating the output of the cerebellar module in early learning impairs further learning and performance (Fig. 7a,b). This is expected as the cerebellar network provides feedback predictions that facilitate cerebral learning. Interestingly, we observe that when the cerebellum is suddenly removed learning becomes worse than the baseline model. This is likely due to the additional time taken to adapt to a new learning trajectory which no longer relies on cerebellar prediction. However, cerebellar lesions performed later in learning do not have an impact

in the simple LD visuomotor task, which is explained by the fact that for this task the model can achieve near-zero error, thus learning signals provided by the cerebellum are no longer needed. However, for all the online tasks we observe that inactivating the cerebellum even at later stages damages learning. In these more realistic tasks the cortical network still relies on the feedback provided by cerebellum as it does not fully learn the task. Our results indicate that lesion studies should reveal a task-dependent nonlinear role of the cerebellum on cerebral learning.

Next, we assess the impact of disrupting cerebellar learning by modelling a complete lesion of our inferior olive-like error module (Methods). This manipulation effectively stops cerebellar learning, thereby impacting on the ability of the cerebellum to provide informative feedback learning signals to the cerebral network which may prevent the cerebral network from learning. For all of the tasks that we model, inactivating cerebellar learning has a strong impact throughout training, making the model return to naive performance (Fig. 7c,d). Thus, simulated "inferior olive" lesions predicts that if the cerebellum cannot learn it would result in a stronger negative impact in task learning than ablating the cerebellum itself. This further suggests that it is critical for the cerebellum to learn rapidly to be able to provide informative predictions.

Cerebro-cerebellar model facilitates learning in a visual-language task

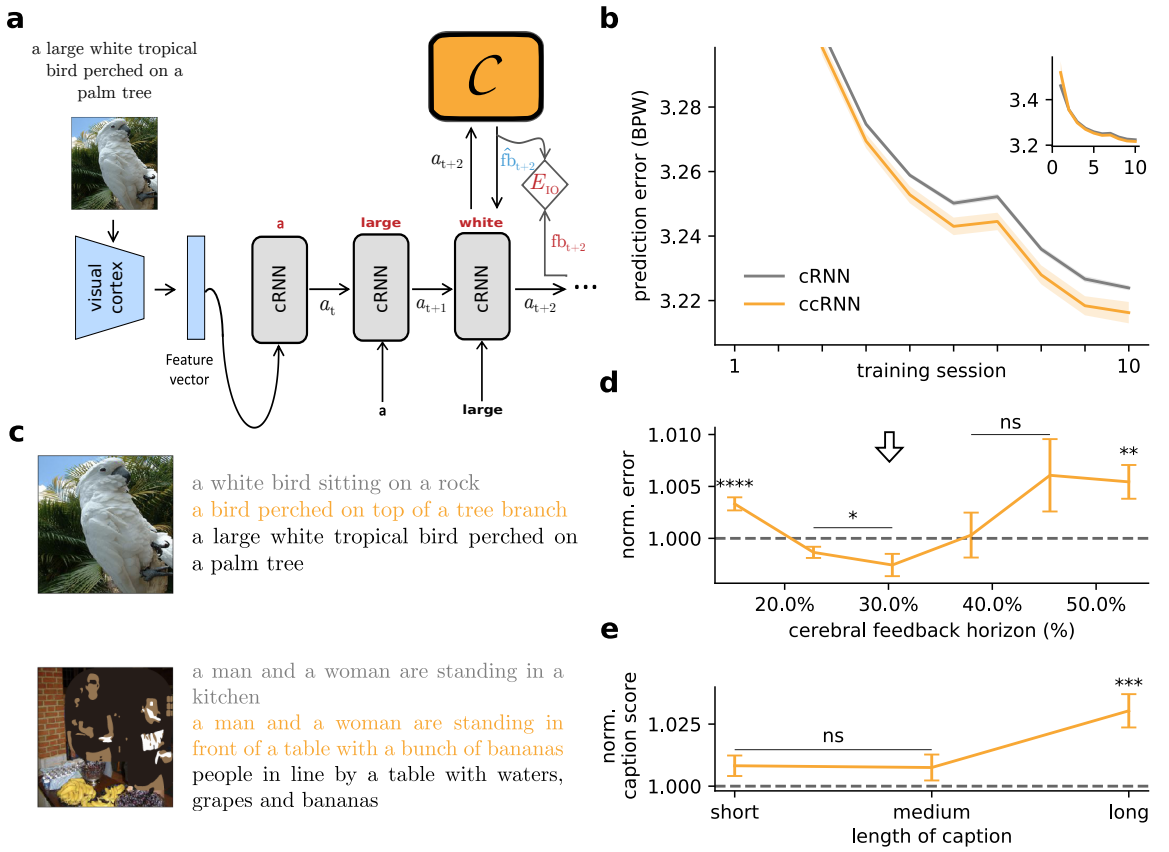


Figure 8. Cerebro-cerebellar model facilitates learning in a visual-language task. (**a**) Schematic of the model used in a visual-language task. The image is first processed by a (pretrained) convolutional neural network modelling the visual cortex. The resulting feature vector is then provided to the cerebral RNN which is trained to predict the next word given the previous words of a provided "gold standard" caption to the image. The cerebellum module $\mathcal{C}$ is only applied to the cRNN. (**b**) Learning curves in bits per word (BPW), lower values indicate better understanding of the language, on validation set for cerebral feedback horizon of four timesteps (inset shows complete learning curve). (**c**) Two example images from the validation set with corresponding model captions and gold standard captions (black). (**d**) Normalised model performance across different degrees of feedback horizon in the cerebral network (p=0.891 (40%), p=0.116 (45%). (**e**) Normalised caption score (Methods) as a function of caption length (p=0.075 (short), p=0.189(medium)). *: p<0.05, **: p<0.01, ***: p<0.001, ****: p<0.0001. Error bars represent mean ± SEM across 10 different initial conditions.

Our framework does not only apply to sensorimotor tasks, but should generalise to virtually any task within the

grasp of current neural networks models. To test the generability of our model and inspired by cognitive tasks in which cerebellar patients have shown deficits[49] we test our models in a caption generation task. In this task the network needs to generate a textual description for a given image. All models have two components: a pretrained convolutional neural network (CNN) to extract a lower dimensional representation of the image, and a cRNN or ccRNN on top which is trained to map the low dimensional visual input to captions that describe the image. (Fig. 8a).

We use a standard machine learning dataset[50] and the networks are trained to predict the next word (Methods). In contrast to the previous tasks here we use a form of unsupervised learning, in which the prediction error module only uses the data itself (i.e. words) to generate teaching signals (Supplementary). We find that ccRNN models can exhibit faster learning (Fig. 8b) (Fig. 8d) and better generalisation[51] (Fig. S9) when in the presence of short cerebral feedback horizons ($\leq$ 40%). All models produce reasonable captions for images unseen during training, but ccRNN models tend to produce captions that better capture the context and semantics of the task (Figs. 8c, S10), consistent with cerebellar deficits[14].

Finally, we use a language metric (SPICE[52]) to measure the quality of the generated captions. These results show that the ccRNN generates richer captions (Fig. 8e) and that it is particularly beneficial for longer captions. This suggests that ccRNN is able to learn richer visuo-language contextual information.

Discussion

Inspired by recent deep learning developments, here we have introduced a systems-level computational model in which cerebellar networks predict cerebral feedback (Fig. 1). In this scheme cerebro-cerebellar loops decouple cerebral cortical networks from future feedback signals. We show that the ccRNN model accelerates learning and improves task behaviour in a range of sensorimotor and cognitive tasks (Figs. 2, 3 and 8). Our results are consistent with observed motor and cognitive deficits in cerebellar patients. Our model makes a number of predictions in terms of (1) task properties (Figs. 4 and 5), (2) cerebro-cerebellar representations and coupling (Figs. 2 and 6), and (3) the differential role of the cerebellum and the inferior olive throughout learning (Fig. 7).

Experimental studies have shown that incomplete or delayed external sensory feedback is important for learning[45,53,54]. Our model proposes that the cerebellum plays an important role in facilitating motor learning when in the presence of incomplete or delayed feedback. Furthermore, our work suggests that cerebro-cerebellar networks are ideally placed to facilitate learning when task feedback is presented intermittently, at medium frequencies with respect to task sequence. Similarly, our results suggest that cerebellum-dependent dysmetria should be more prevalent for tasks with intermediate to long inter-feedback intervals. Although there is a wide range of studies investigating the role of external sensory feedback in learning[53,55] and the precise timing of feedback is known to be important for cerebellar function[10,56], it remains to be tested what are the optimal properties of task feedback for learning. Taken together, we suggest cerebellar-mediated feedback predictions to be particularly important for temporally challenging tasks with sparse feedback.

Our representational analyses demonstrate that the cerebellum develops task-specific representations. Recent fMRI studies have observed that different regions of the cerebellum encodes task-specific representations for different domains[23,57]. Similarly, our model predicts the need for different cerebellar modules to provide feedback estimations to the cerebral cortex for specific task domains. We have also studied the level of coupling between cerebellar and cerebral neural activity. Our results demonstrate an initial rise in correlations which coincides with steep periods of learning followed by a general decay in the coupling during the remaining periods of learning. This general decay in coupling is also reflected in our simulated cerebellar lesions which echo the existing literature in that after a task is consolidated in the cerebrum it becomes less cerebellar-dependent[58,59].

In line with previous theoretical accounts[6,7,9] we suggest that the cerebellar error function is computed by the *inferior olive*, which drives learning in the cerebellum via the climbing fibres. This cerebellar error function is a combination of true sensory feedback and self-predicted (bootstrapped) error signals (Fig. 1b), which is analogous to the bootstrapping principles commonly used in reinforcement learning[60]. The use of self-predictions in the cerebellum suggests the existence of different forms of feedback to the inferior olive from potentially multiple cerebellar modules, consistent with cerebellar-inferior olive connectivity[61]. Moreover, when ablating the inferior olive lesions we show that task performance become severely impaired. This is due to the cerebellum being unable to learn, thereby providing outdated feedback signals back to the cerebral cortex. These results suggest non-trivial consequences of lesions for cerebro-cerebellar interactions.

While our model is consistent with experimental observations, there are several biological features that we have not considered. In particular, experimental studies suggest that the cerebellum can influence cerebral learning processes via its projections to the thalamus [62–65]. This is in line with ccRNN where the cerebellum predicts feedback signals that contribute directly to cerebral learning. However, we have assumed direct long-range projections with the cerebral cortex whereas in biology these projections are mediated through the thalamus and pons. It is possible that both structures may provide bottlenecks that filter out non-relevant information, such as poor estimated feedback (Figs. 2d, 3d) that would impair cerebral learning. In addition, cerebellar-thalamic-cerebral projections are known to target distal dendrites of pyramidal cells [66,67], which have been proposed to encode feedback error signals by a number of models as used by deep learning models [68,69]. These dendritic-encoded error signals are akin to the gradient descent errors that we use to model cortical feedback signals. In future work it would be of interest to combine our work with these biologically plausible gradient descent models.

Throughout this paper we have assumed the existence of cerebral prediction error modules, which compare the output of a given cerebral area with a desired task output to generate a feedback teaching signal for the cerebral cortex. There is evidence of prediction errors across different brain areas, for example sensorimotor prediction errors in the neocortex [70,71] or reward prediction errors in the VTA [1,72]. For simplicity, here we have focused on supervised (Figs. 2,3) and unsupervised (Fig. 8) prediction errors, but these can in principle be readily replaced by reward-based prediction errors [1,73]. This would predict reward-specific encodings in the cerebellum as observed recently [74–76]. Indeed, our model is of particular relevance to reinforcement learning due to prevalence of sparse and delayed rewards (Fig. 4).

Finally, our model shares common features with classical internal models of the cerebellum ([6,7]; Table S1). In the forward model of sensorimotor control, the cerebellum receives an efferent copy of the motor commands and the respective external sensory feedback [8,77]. With these two input streams the forward model learns to predict the sensory consequences of motor commands. We and others have argued that a similar predictive model can in principle be applied to higher order brain regions such as the prefrontal cortex and the temporo-parietal cortex which are involved in planning of cognitive behaviour and decision making [16,17,24,26] (Fig. 1a). In line with forward models the cerebellar module of ccRNN receives an efferent copy of the cerebral neural activity and cerebral feedback. Given these signals the cerebellum learns to predict future cerebral feedback.

Overall, our work offers a novel theoretical framework with which to study cerebro-cerebellar interactions, being consistent with experimental observations while making a large number of testable predictions across multiple levels of interrogation.

Acknowledgements

We would like to thank the Neural & Machine Learning group, Paul Anastasiades, Paul Dodson, Conor Houghton, Laurence Aitchison, Cian O'Donnell, James M. Shine, Max Jaderberg, Nadia Cerminara and Jasmine Pickford for useful feedback. We would also like to thank Samia Mohinta and Milton Llera Montero for help with model analysis and training. JP was funded by a EPSRC Doctoral Training Partnership award (EP/R513179/1) and EB by the Wellcome Trust (220101/Z/20/Z). This work made use of the HPC system Blue Pebble at the University of Bristol, UK.

References

[1] W. Schultz, P. Dayan, and P. R. Montague. A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599, 1997.

[2] C. D. B. Luft. Learning from feedback: The neural mechanisms of feedback processing facilitating better performance. *Behavioural brain research*, 261:356–368, 2014.

[3] M. H. Herzog and M. Fahle. The role of feedback in learning a vernier discrimination task. *Vision research*, 37(15):2133–2141, 1997.

[4] D. M. Wolpert and Z. Ghahramani. Computational principles of movement neuroscience. *Nature neuroscience*, 3(11):1212–1217, 2000.

[5] D. M. Wolpert, R. C. Miall, and M. Kawato. Internal models in the cerebellum. *Trends in cognitive sciences*, 2(9):338–347, 1998.

[6] D. Marr. A theory of cerebellar cortex. *The Journal of Physiology*, 202(2):437–470, jun 1969. ISSN 00223751. doi: 10.1113/jphysiol.1969.sp008820. URL http://doi.wiley.com/10.1113/jphysiol.1969.sp008820.

[7] J. S. Albus. A theory of cerebellar function. *Mathematical Biosciences*, 10(1):25–61, 1971. ISSN 0025-5564. doi: 10.1016/0025-5564(71)90051-4.

[8] R. C. Miall, D. J. Weir, D. M. Wolpert, and J. F. Stein. Is the cerebellum a smith predictor? *Journal of Motor Behavior*, 25(3):203–216, 1993. ISSN 19401027. doi: 10.1080/00222895.1993.9942050.

[9] J. L. Raymond and J. F. Medina. Computational principles of supervised learning in the cerebellum. *Annual review of neuroscience*, 41:233–253, 2018.

[10] N. L. Cerminara, R. Apps, and D. E. Marple-Horvat. An internal model of a moving visual target in the lateral cerebellum. *The Journal of physiology*, 587(2):429–442, 2009.

[11] J. N. Sanes, B. Dimitrov, and M. Hallett. Motor learning in patients with cerebellar dysfunction. *Brain*, 113(1):103–120, 1990.

[12] J. A. Fiez, S. E. Petersen, M. K. Cheney, and M. E. Raichle. Impaired non-motor learning and error detection associated with cerebellar damage: A single case study. *Brain*, 115(1):155–178, 1992.

[13] N. Rahmati, C. B. Owens, L. W. Bosman, J. K. Spanke, S. Lindeman, W. Gong, J.-W. Potters, V. Romano, K. Voges, L. Moscato, et al. Cerebellar potentiation and learning a whisker-based object localization task with a time response window. *Journal of Neuroscience*, 34(5):1949–1962, 2014.

[14] X. Guell, F. Hoche, and J. D. Schmahmann. Metalinguistic Deficits in Patients with Cerebellar Dysfunction: Empirical Support for the Dysmetria of Thought Theory. *Cerebellum*, 14(1):50–58, December 2015. ISSN 14734230. doi: 10.1007/s12311-014-0630-z.

[15] T. M. Locke, M. E. Soden, S. M. Miller, A. Hunker, C. Knakal, J. A. Licholai, K. S. Dhillon, C. D. Keene, L. S. Zweifel, and E. S. Carlson. Dopamine D1 Receptor–Positive Neurons in the Lateral Nucleus of the Cerebellum Contribute to Cognitive Behavior. *Biological Psychiatry*, 84(6):401–412, 2018. ISSN 18732402. doi: 10.1016/j.biopsych.2018.01.019.

[16] J. A. Brissenden and D. C. Somers. Cortico–cerebellar networks for visual attention and working memory. *Current Opinion in Psychology*, 2019. ISSN 2352250X. doi: 10.1016/j.copsyc.2019.05.003.

[17] J. D. Schmahmann, X. Guell, C. J. Stoodley, and M. A. Halko. The Theory and Neuroscience of Cerebellar Cognition. *Annual Review of Neuroscience*, 42(1):337–364, July 2019. ISSN 0147-006X. doi: 10.1146/annurev-neuro-070918-050258.

[18] X. Guell, J. D. Gabrieli, and J. D. Schmahmann. Triple representation of language, working memory, social and emotion processing in the cerebellum: Convergent evidence from task and seed-based resting-state fMRI analyses in a single large cohort. *NeuroImage*, 172:437–449, May 2018. ISSN 10959572. doi: 10.1016/j.neuroimage.2018.01.082.

[19] B. Deverett, M. Kislin, D. W. Tank, S. Samuel, and H. Wang. Cerebellar disruption impairs working memory during evidence accumulation. *bioRxiv*, page 521849, 2019.

[20] S. C. Baker, R. D. Rogers, A. M. Owen, C. D. Frith, R. J. Dolan, R. S. J. Frackowiak, and T. W. Robbins. Neural systems engaged by planning: a PET study of the Tower of London task. *Neuropsychologia*, 34(6):515–526, 1996.

[21] R. M. Kelly and P. L. Strick. Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *Journal of neuroscience*, 23(23):8432–8444, 2003.

[22] P. L. Strick, R. P. Dum, and J. A. Fiez. Cerebellum and nonmotor function. *Annual review of neuroscience*, 32:413–434, 2009.

[23] M. King, C. R. Hernandez-Castillo, R. A. Poldrack, R. B. Ivry, and J. Diedrichsen. Functional boundaries in the human cerebellum revealed by a multi-domain task battery. *Nature Neuroscience*, 22(8):1371–1378, August 2019.

[24] M. Ito. Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9(4):304–313, apr 2008. ISSN 1471-003X. doi: 10.1038/nrn2332. URL www.nature.com/reviews/neurohttp://www.nature.com/articles/nrn2332.

[25] M. J. Wagner, T. H. Kim, J. Kadmon, N. D. Nguyen, S. Ganguli, M. J. Schnitzer, and L. Luo. Shared Cortex-Cerebellum Dynamics in the Execution and Learning of a Motor Task. *Cell*, 177(3):669–682.e24, 2019. ISSN 10974172. doi: 10.1016/j.cell.2019.02.019. URL https://doi.org/10.1016/j.cell.2019.02.019.

[26] M. J. Wagner and L. Luo. Neocortex–Cerebellum Circuits for Cognitive Processing. *Trends in Neurosciences*, 43(1):42–54, 2020. ISSN 1878108X. doi: 10.1016/j.tins.2019.11.002. URL https://doi.org/10.1016/j.tins.2019.11.002.

[27] V. Mante, D. Sussillo, K. V. Shenoy, and W. T. Newsome. Context-dependent computation by recurrent dynamics in prefrontal cortex. *nature*, 503(7474):78–84, 2013.

[28] H. F. Song, G. R. Yang, and X.-J. Wang. Training excitatory-inhibitory recurrent neural networks for cognitive tasks: a simple and flexible framework. *PLoS computational biology*, 12(2):e1004792, 2016.

[29] K. Rajan, C. D. Harvey, and D. W. Tank. Recurrent network models of sequence generation and memory. *Neuron*, 90(1):128–142, 2016.

[30] R. Laje and D. V. Buonomano. Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nature neuroscience*, 16(7):925–933, 2013.

[31] P. A. Butcher, R. B. Ivry, S. H. Kuo, D. Rydz, J. W. Krakauer, and J. A. Taylor. The cerebellum does more than sensory prediction error-based learning in sensorimotor adaptation tasks. *Journal of Neurophysiology*, 118(3):1622–1636, 2017. ISSN 15221598. doi: 10.1152/jn.00451.2017.

[32] A. Nashef, O. Cohen, R. Harel, Z. Israel, and Y. Prut. Reversible Block of Cerebellar Outflow Reveals Cortical Circuitry for Motor Coordination. *Cell Reports*, 27(9):2608–2619.e4, 2019. ISSN 22111247. doi: 10.1016/j.celrep.2019.04.100. URL https://doi.org/10.1016/j.celrep.2019.04.100.

[33] M. Jaderberg, W. M. Czarnecki, S. Osindero, O. Vinyals, A. Graves, D. Silver, and K. Kavukcuoglu. Decoupled neural interfaces using synthetic gradients. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1627–1635. JMLR. org, 2017.

[34] S. Herculano-Houzel. The human brain in numbers: a linearly scaled-up primate brain. *Frontiers in Human Neuroscience*, 3:31, 2009. ISSN 1662-5161. doi: 10.3389/neuro.09.031.2009. URL https://www.frontiersin.org/article/10.3389/neuro.09.031.2009.

[35] J. Diedrichsen, M. King, C. Hernandez-Castillo, M. Sereno, and R. B. Ivry. Universal transform or multiple functionality? understanding the contribution of the human cerebellum across task domains. *Neuron*, 102(5):918–928, 2019.

[36] Y.-w. Tseng, J. Diedrichsen, J. W. Krakauer, R. Shadmehr, and A. J. Bastian. Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of neurophysiology*, 98(1):54–62, 2007.

[37] M. L. Streng, L. S. Popa, and T. J. Ebner. Modulation of sensory prediction error in purkinje cells during visual feedback manipulations. *Nature communications*, 9(1):1–12, 2018.

[38] J. Hore, B. Wild, and H. Diener. Cerebellar dysmetria at the elbow, wrist, and fingers. *Journal of neurophysiology*, 65(3):563–571, 1991.

[39] S. E. Criscimagna-Hemminger, A. J. Bastian, and R. Shadmehr. Size of error affects cerebellar contributions to motor learning. *Journal of neurophysiology*, 103(4):2275–2284, 2010.

[40] Y. LeCun, C. Cortes, and C. Burges. Mnist handwritten digit database, 2010.

[41] J. D. Schmahmann. Disorders of the cerebellum: ataxia, dysmetria of thought, and the cerebellar cognitive affective syndrome. *The Journal of neuropsychiatry and clinical neurosciences*, 16(3):367–378, 2004.

[42] Z. Gao, C. Davis, A. M. Thomas, M. N. Economo, A. M. Abrego, K. Svoboda, C. I. De Zeeuw, and N. Li. A cortico-cerebellar loop for motor planning. *Nature*, 563(7729):113–116, 2018.

[43] M. King, C. R. Hernandez-Castillo, R. A. Poldrack, R. B. Ivry, and J. Diedrichsen. Functional boundaries in the human cerebellum revealed by a multi-domain task battery. *Nature neuroscience*, 22:1371–1378, 2019.

[44] S. Kitazawa, T. Kohno, and T. Uka. Effects of delayed visual information on the rate and amount of prism adaptation in the human. *Journal of Neuroscience*, 15(11):7644–7652, 1995.

[45] T. Ikegami, M. Hirashima, R. Osu, and D. Nozaki. Intermittent visual feedback can boost motor learning of rhythmic movements: evidence for error feedback beyond cycles. *Journal of Neuroscience*, 32(2):653–657, 2012.

[46] N. Levy, A. Pressman, F. A. Mussa-Ivaldi, and A. Karniel. Adaptation to delayed force perturbations in reaching movements. *PLoS one*, 5(8):e12128, 2010.

[47] M. Synofzik, A. Lindner, and P. Thier. The Cerebellum Updates Predictions about the Visual Consequences of One's Behavior. *Current Biology*, 18(11):814–818, 2008. ISSN 09609822. doi: 10.1016/j.cub.2008.04.071.

[48] C. Batcho, M. Gagné, L. Bouyer, J. Roy, and C. Mercier. Impact of online visual feedback on motor acquisition and retention when learning to reach in a force field. *Neuroscience*, 337:267–275, 2016.

[49] A. L. Gebhart, S. E. Petersen, and W. T. Thach. Role of the posterolateral cerebellum in language. *Annals of the New York Academy of Sciences*, 978(1):318–333, 2002.

[50] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015. doi: 10.1007/s11263-015-0816-y.

[51] W. M. Czarnecki, G. Swirszcz, M. Jaderberg, S. Osindero, O. Vinyals, and K. Kavukcuoglu. Understanding synthetic gradients and decoupled neural interfaces. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 904–912. JMLR. org, 2017.

[52] P. Anderson, B. Fernando, M. Johnson, and S. Gould. Spice: Semantic propositional image caption evaluation. In *European Conference on Computer Vision*, pages 382–398. Springer, 2016.

[53] T. Honda, M. Hirashima, and D. Nozaki. Adaptation to visual feedback delay influences visuomotor learning. *PLoS one*, 7(5): e37900, 2012.

[54] S. H. Park, A. Casamento-Moran, M. L. Singer, A. E. Ernster, B. Yacoubi, I. A. Humbert, and E. A. Christou. Integration of visual feedback and motor learning: corticospinal vs. corticobulbar pathway. *Human movement science*, 58:88–96, 2018.

[55] A. J. M. Foulkes and R. C. Miall. Adaptation to visual feedback delays in a human manual tracking task. *Experimental brain research*, 131(1):101–110, 2000.

[56] H. Beppu, M. Suda, and R. Tanaka. Analysis of cerebellar motor disorders by visually guided elbow tracking movement. *Brain*, 107(3):787–809, 1984.

[57] R. Ashida, N. L. Cerminara, R. J. Edwards, R. Apps, and J. C. W. Brooks. Sensorimotor, language, and working memory representation within the human cerebellum. *Human Brain Mapping*, 40(16):4732–4747, July 2019.

[58] J. Doyon, V. Penhune, and L. G. Ungerleider. Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia*, 41(3):252–262, 2003.

[59] E. Galliano, Z. Gao, M. Schonewille, B. Todorov, E. Simons, A. S. Pop, E. D'Angelo, A. M. J. M. Van Den Maagdenberg, F. E. Hoebeek, and C. I. De Zeeuw. Silencing the majority of cerebellar granule cells uncovers their essential role in motor learning and consolidation. *Cell reports*, 3(4):1239–1251, 2013.

[60] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[61] F. Bengtsson and G. Hesslow. Cerebellar control of the inferior olive. *The cerebellum*, 5(1):7–14, 2006.

[62] S. E. Hua and J. C. Houk. Cerebellar guidance of premotor network development and sensorimotor learning. *Learning & Memory*, 4(1):63–76, 1997.

[63] V. B. Penhune and J. Doyon. Cerebellum and m1 interaction during early learning of timed motor sequences. *Neuroimage*, 26 (3):801–812, 2005.

[64] A. Kishore, S. Meunier, and T. Popa. Cerebellar influence on motor cortex plasticity: behavioral implications for parkinson's disease. *Frontiers in neurology*, 5:68, 2014.

[65] Y. H. Tanaka, Y. R. Tanaka, M. Kondo, S.-I. Terada, Y. Kawaguchi, and M. Matsuzaki. Thalamocortical axonal activity in motor cortex exhibits layer-specific dynamics during motor learning. *Neuron*, 100(1):244–258, 2018.

[66] K. Guo, N. Yamawaki, K. Svoboda, and G. M. Shepherd. Anterolateral motor cortex connects with a medial subdivision of ventromedial thalamus through cell type-specific circuits, forming an excitatory thalamo-cortico-thalamic loop via layer 1 apical tuft dendrites of layer 5b pyramidal tract type neurons. *Journal of Neuroscience*, 38(41):8787–8797, 2018.

[67] P. G. Anastasiades, D. P. Collins, and A. G. Carter. Mediodorsal and ventromedial thalamus engage distinct l1 circuits in the prefrontal cortex. *Neuron*, 109(2):314–330, 2021.

[68] J. Sacramento, R. P. Costa, Y. Bengio, and W. Senn. Dendritic cortical microcircuits approximate the backpropagation algorithm. In *Advances in Neural Information Processing Systems*, pages 8721–8732, 2018.

[69] A. Payeur, J. Guerguiev, F. Zenke, B. A. Richards, and R. Naud. Burst-dependent synaptic plasticity can coordinate learning in hierarchical circuits. *Nature neuroscience*, pages 1–10, 2021.

[70] A. Attinger, B. Wang, and G. B. Keller. Visuomotor Coupling Shapes the Functional Development of Mouse Visual Cortex. *Cell*, 169(7):1291–1302.e14, June 2017. ISSN 0092-8674. doi: 10.1016/j.cell.2017.05.023.

[71] R. Jordan and G. B. Keller. Opposing Influence of Top-down and Bottom-up Input on Excitatory Layer 2/3 Neurons in Mouse Primary Visual Cortex. *Neuron*, 108(6):1194–1206.e5, December 2020. ISSN 0896-6273. doi: 10.1016/j.neuron.2020.09.024.

[72] R. Keiflin and P. H. Janak. Dopamine prediction errors in reward learning and addiction: From theory to neural circuitry. *Neuron*, 88(2):247–263, 2015.

[73] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, February 2015. ISSN 1476-4687. doi: 10.1038/nature14236.

[74] M. J. Wagner, T. H. Kim, J. Savall, M. J. Schnitzer, and L. Luo. Cerebellar granule cells encode the expectation of reward. *Nature*, 544(7648):96–100, 2017.

[75] I. Carta, C. H. Chen, A. L. Schott, S. Dorizan, and K. Khodakhah. Cerebellar modulation of the reward circuitry and social behavior. *Science*, 363(6424), 2019.

[76] N. Sendhilnathan, A. E. Ipata, and M. E. Goldberg. Neural correlates of reinforcement learning in mid-lateral cerebellum. *Neuron*, 2020.

[77] M. Ito. Neurophysiological aspects of the cerebellar motor control system. *International journal of neurology*, 7(2):162–176, 1970. ISSN 0020-7446. URL http://europepmc.org/abstract/MED/5499516.

[78] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[79] R. P. Costa, I. A. Assael, B. Shillingford, N. de Freitas, and T. Vogels. Cortical microcircuits as gated-recurrent neural networks. In *Advances in neural information processing systems*, pages 272–283, 2017.

[80] H. Tanaka, T. Ishikawa, J. Lee, and S. Kakei. The Cerebro-Cerebellum as a Locus of Forward Model: A Review. *Frontiers in Systems Neuroscience*, 14:19, 2020. ISSN 1662-5137. doi: 10.3389/fnsys.2020.00019. URL https://www.frontiersin.org/article/10.3389/fnsys.2020.00019.

[81] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[82] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034, 2015.

[83] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan. Show and tell: A neural image caption generator. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3156–3164, 2015.

[84] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[85] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.

[86] D. Kobak, W. Brendel, C. Constantinidis, C. E. Feierstein, A. Kepecs, Z. F. Mainen, X.-L. Qi, R. Romo, N. Uchida, and C. K. Machens. Demixed principal component analysis of neural population data. *eLife*, 5, 2016. ISSN 2050-084X. doi: 10.7554/eLife.10989.

Methods

In all our experiments we model a cerebral area A as a long short-term memory recurrent neural network (LSTM)[78] with parameters θ which has recently been mapped onto cortical microcircuits[79]. A (trained) linear readout is attached to the LSTM output states which provides the final model output to which a supervised error module E^{task}, which below we refer to as E.

In the cerebro-cerebellar RNN model (ccRNN) we attach a feedforward cerebellar module C with independent parameters Ψ to the RNN with reciprocal connections (Fig. 1). The cerebellar module is equivalent to the "synthesiser" as used by Jaderberg et al.[33] in the backward case. That is, the cerebellar module receives a copy of the RNN activity a_t (both cell and output LSTM states) and sends back a prediction of the future feedback (or error gradients) with respect to that activity, $C(a_t)$.

To generate the desired cerebral temporal feedback (error gradients) we use backpropagation through time (BPTT). To highlight the link between BPTT-derived feedback and the cerebellar predicted feedback we start out from first principles closely following Jaderberg et al.[33]. BPTT is used as the standard solution for updating parameters θ in deep learning. In an ideal world one would have access to all error signals within a task of length T, $\sum_t^T E_t$, and derive the resulting parameter updates as $\theta \leftarrow \theta - \alpha\Delta\theta$ where $\Delta\theta = \sum_{t'=t}^{T} \frac{\partial E_{t'}}{\partial \theta}$, but this is impractical as it would require information about all possible future error signals. Instead, in practice BPTT over a time horizon K is commonly used (Fig. 1)

$$
\begin{aligned}
\sum_{t'=t}^{T} \frac{\partial E_{t'}}{\partial \theta} &= \left(\sum_{t'=t}^{K} \frac{\partial E_{t'}}{\partial a_t} \right) \frac{\partial a_t}{\partial \theta} + \left(\sum_{t'=K+1}^{T} \frac{\partial E_{t'}}{\partial a_K} \right) \frac{\partial a_K}{\partial \theta} \\
\sum_{t'=t}^{T} \frac{\partial E_{t'}}{\partial \theta} &\approx \left(fb_{t'} + \hat{fb}_{t'>K} \frac{\partial a_K}{\partial a_t} \right) \frac{\partial a_t}{\partial \theta} \\
\Delta\theta &\approx \left(\underbrace{fb_t}_{\text{cerebral feedback}} + \underbrace{C(a_K)}_{\text{cerebellar feedback}} \frac{\partial a_K}{\partial a_t} \right) \frac{\partial a_t}{\partial \theta}
\end{aligned}
\tag{1}
$$

where $fb_{t'}$ denotes the cerebral feedback and $C(a_K)$ cerebellar predictions of *future feedback* and $\frac{\partial a_K}{\partial a_t}$ represents the temporal changes in cerebral activity. These equations help make clear the distinction between cerebral feedback modelling feedback within current horizon K and cerebellar feedback predicting future horizons. Note that if we set $C(a_{t'}) = 0$ then we simply have standard truncated-BPTT over a time horizon K, as commonly used in deep learning.

A key consequence of the cerebellum predicting future feedback is that *strong* or *long* feedback signals (i.e. $T \gg 0$) are no longer necessary, thus decoupling learning in the cerebral network from future feedback signals. For this reason we focus on weak forms of BPTT with relatively small temporal horizons, in which we model only K time steps of feedback into the past from an error signal E, this is known as truncated BPTT. In our experiments the size of K - which we report as a percentage of the task length (cerebral temporal gradient) - varies but is generally small. For example, for the simple line drawing task we used a one-step BPTT (i.e. $K = 1$; Fig. 2). Note in the main text as we focus on describing a simpler case of $K = 1$ (as used in the simple line drawing task) we use $C(a_t)$ to refer to the cerebellar feedback prediction from the end of the current horizon, i.e. $C(a_t) = C(a_K)\frac{\partial a_K}{\partial a_t}$.

Cerebellar learning

The cerebellar parameters Ψ are themselves learnt but to optimise a distinct, specialised error E^{IO} which we posit to be computed at the *inferior olive*, the classical teacher of the cerebellar cortex[6,7]. This is defined by the difference between cerebellar output and a target feedback signal $\bar{fb}_t$, i.e. $E_t^{\text{IO}} = ||C(a_t) - \bar{fb}_t||$. Similar to the cerebral network we update cerebellar parameters using gradient descent: $\Psi \leftarrow \Psi - \alpha^{IO}\Delta\Psi$, where $\Delta\Psi = \frac{\partial E^{\text{IO}}}{\partial \Psi}$.

Ideally we would simply set the target feedback as the true (desired) cerebral feedback. However, this would require an arbitrary long number of steps of true cerebellar feedback, exactly what we propose that is not required with a cerebellar network. How should the cerebellum itself learn about the future feedback? One elegant solution, which we take from Jaderberg et al.[33], is to combine the currently available error with bootstrapped future cerebellar predictions (i.e. self-predictions). Formally, using the same notation as equation 1, the trained target for $C(a_T)$ is

$$
\bar{fb}_T = \frac{\partial E_{\leq 2T}}{\partial a_T} + C(a_{2T})\frac{\partial a_{2T}}{\partial a_T} \tag{2}
$$

Note the resemblance of equation 2 to equation 1: in each case we consider a mixture of nearby "cerebral" error signals beyond which we rely on cerebellar prediction. It is also useful to compare equation 2 with standard reinforcement learning rules (e.g. temporal difference learning algorithm) which rely on similar bootstrapping principles [60].

Other biological mappings of our framework

Here we describe other possible mappings between the proposed framework (cerebellum as a decoupling machine) and forward and feedback processing in the cerebral cortex.

Cerebellum as a spatial feedback decoupler

Our paper focuses on temporal problems being solved by a cerebral area modelled as a recurrent neural network (RNN) to which a cerebellar network provides predictions of the future errors/feedback with respect to that area. An analogous biologically relevant system also arises, however, when one considers cerebral processing in space using feedforward computations involving several distinct regions (Fig. S1).

This setup - where the "main" (cerebral) network is a feedforward composition of multiple brain regions - was also considered in Jaderberg et al.. Now, as opposed to predicting errors which occur strictly at later points in time, the role of the cerebellar network is to predict errors which occur in later brain regions. The result is that an earlier region has access to its feedback (predicted by the cerebellum) without the need to wait for the later forward/back propagation of spatial activity. Formally, if we assume cerebral processing as a sequence $\{a_i\}_{i=1}^{N}$ of feedforward computations: $A(x) = (a_N \circ a_{N-1} \circ \cdots \circ a_1)(x)$ which defines a final error function $E\big(A(x)\big)$, then the cerebellar network can provide predicted feedback at a given brain area as soon as its activities are computed: $\mathcal{C}(a_i) := \hat{fb}_i = \frac{\partial \hat{E}}{\partial a_i} \approx \frac{\partial E}{\partial a_i}$.

This perspective would effectly feedback processing across the brain. This interpretation of the model is consistent with cerebellar-thalamo-cerebral projections targeting distal dendrites, which have been proposed as the site of error or feedback encoding which underlie efficient learning [68,69].

Cerebellum as a forward decoupler

In classical cerebellar theory, the complement to the forward model hypothesis is the inverse model, in which the cerebellum predicts motor commands [5], or even implicit mental predictions to solve a problem [24], directly . Again we can consider this under the proposed framework, but now using its *forward* prediction version.

In this case the role of the cerebellum is not to predict future feedback activity, but the feedforward activity itself, i.e., $\mathcal{C}(a_i) = \hat{a}_j$ for some later region $j > i$. $\hat{a}_j$ is fed as a replacement to region j, making it forward decoupled from a potentially slower intermediate processing $a_j \circ a_{j-1} \circ \cdots \circ a_{i+1}$.

Functionally this would provide the organism with fast inputs (e.g. motor commands or potential mental solutions) without the need for potentially slower cerebral processing (Fig. S1b). We also point out the relevance of direct predictions of later activity in the temporal case, where the cerebellum strictly predicts motor activity at later timesteps, as suggested in [80]. A broad comparison between this framework and the cerebellar internal model hypothesis is shown in Table S1.

Experimental details

To reduce learning instability we scale the cerebellar predicted feedback (Eq. 1) by 0.1 [33]. Both cerebral and cerebellar parameters are optimised using the feedback described above together with ADAM for overall learning efficiency [81]. Training the model involves iterating over training sessions for a given dataset, which is split into batches. For better learning stability model parameters were updated at the end of each batch.

In each experiment all initial RNN parameters are drawn from an uniform distribution $\mathcal{U}(-\frac{1}{\sqrt{n_{\text{RNN}}}}, \frac{1}{\sqrt{n_{\text{RNN}}}})$, where n_{RNN} is the number of RNN units. The weights of the readout network and the feedforward weights of the cerebellar network (other than the final layer) are initialised according to $\mathcal{U}(-b_k, b_k)$ where b_k denotes the "kaiming bound" as described by He et al. [82] (slope $a = \sqrt{5}$), and the biases are draw from $\mathcal{U}(-\frac{1}{\sqrt{n_{\text{in}}}}, \frac{1}{\sqrt{n_{\text{in}}}})$, where n_{in} denotes the input size of the layer. The last layer (both weights and bias) of the cerebellar network is zero-initialised, so that the estimated feedback at the start are zero [33].

During learning, we employ truncated BPTT as follows. Given an input sequence of N timesteps $x_1, x_2, \ldots, x_N$ and an temporal horizon K, we divide the sequence into K sized truncations. In other words, the sequence is now made

up truncations of $(x_1, \dots, x_K), \dots, (x_{(m-1)K+1}, \dots, x_{mK}), (x_{N-r}, \dots, x_N)$, where $N = mT + r$ for positive integers m, r with $0 \leq r < K$. Note that, along with the value K, how well the sequence is divided into truncations (i.e. values m, r) can itself influence learning (e.g. Fig. 3d).

In the all visuomotor tasks, to test the effect of predicted feedback against the availability of task feedback signals which occur at any timestep where an external teaching signal is provided, we vary the *external feedback interval*. Given feedback interval n, the target is only available every n timesteps. This is analogous to the rate at which one receives sensory information whilst performing a task (e.g. drawing freehand).

In general, (standard) hyperparameters were selected by hand after a few trial runs. We used the PyTorch library for all neural network models. Our implementation is based on that of github.com/koz4k/dni-pytorch. The code used for our experiments is available at https://github.com/neuralml/ccDNI.

Delta and normalised error

To calculate the delta and normalised error with respect to a given model we take the difference or ratio of total errors during learning (all training sessions). For example, the normalised error of ccRNN with respect to cRNN is $\frac{\text{error(ccRNN)}}{\text{error(cRNN)}}$. Note that in the ablation case we compare against an "healthy" ccRNN and only consider the respective errors post-ablation. e.g. the normalised error for a model with cerebellar ablation at session 50 is $\frac{\text{error(ablated)}_{>50}}{\text{error(ccRNN)}_{>50}}$.

Cerebro-cerebellar coupling

To analyse how the coupling between the cerebral and cerebellar networks changes over learning we consider the (absolute) Pearson correlation between a given cerebral (LSTM) unit and a given unit in the cerebellar hidden (granular) layer over different bins during training. Values given are the average correlation over all RNN/cerebellar unit pairs.

Computing details

All experiments were conducted on the BluePebble super computer at the university of Bristol; mostly on GPUs (GeForce RTX 2080 Ti) and some on CPUs. We estimate the total compute time (including unreported results) to be in the order of $\sim$ 2000 hours.

Simple line drawing visuomotor task

In the line drawing task, an LSTM network receives a discrete input cue which signals the network to either (1.) stay at zero or (2.) draw a line in 2D space over a period of 10 timesteps. Here we set 6 distinct non-zero input-target pairs $\{(x_i, y_i)\}_{i=1}^6$, where each input x_i is a (one dimensional) integer $\in \{\pm 1, \pm 2, \pm 3\}$, $y_1 = 0$ throughout, and the remaining targets $\{y_i\}_{i=2}^6$ are lines whose end points lie equidistantly on a circle centred on the origin with radius 10. To make the task more realistic we also consider a 7th target in which the network must remain quiet at the centre of the drawing screen, which models periods in which the animal is not actively performing the task. Once an input cue is received at timestep t_0, the model receives no new information (i.e. all future input is set to zero). The model is trained to minimise the mean squared error (MSE) between its output and the cue-based target.

The cerebral network is modelled by one hidden layer of 50 LSTM units and the cerebellar network by one hidden layer of 400 neurons. The learning rate is set to 0.001. Each epoch comprises of 20 batches with 50 randomised examples. Unless explicitly stated we use a truncation size of $K = 1$ which covers 10% of the total task duration. Model results are averaged over 10 random seeds (with error bars), where each seed determines the initial weights of the network.

Online visuomotor tasks

For each online visuomotor task (Fig. 3) we use a standard dataset of handwritten digits (MNIST dataset). The model receives the same temporal input input, and the tasks are only differentiated by the desired model output. Given a 28×28 handwritten digit as input, at timestep i the model receives the pixels from row i of the image, so that the input is of dimension 28 and is presented over 28 timesteps.

In each case we have one hidden layer of 30 LSTM units in the main model and one hidden layer of 300 hidden units in the feedforward cerebellar network. Data was presented in batches of 50 with a learning rate of 0.0001.

Training and validation data was assigned a $4:1$ split, containing 48000 and 12000 distinct image/number pairs respectively. Unless explicitly stated, the truncation value was $K = 3$ which is $\sim 10\%$ of the task duration. Model results are presented over 10 random seeds.

Online line drawing visuomotor task

In this variant each number 0-9 MNIST image is allocated an associated xy position on the edge of a circle centred at 0 with radius 10, and must follow a line of equally spaced points towards that position (Fig. 3a, left). With the model output being a vector of size 2, the training loss is defined at the end by the mean squared error (MSE) between the output of the model and the points forming the target line.

Online digit drawing visuomotor task

Like the online line drawing task, in this variant the model outputs a sequence of 2D coordinates corresponding to the input image. The target sequence however is now of a highly non-linear form, and in this case is a template of the number given as input (Fig. 3a, middle). The model is then trained at to minimise the MSE between the model output and that target shape.

For each digit, the corresponding target drawing lies in $[0, 1] \times [0, 1]$, such that the gap between each successive point is equivalent. All model drawings begin in the top left corner (except for digit 1 which begins below-right). MSE scores are reported as 100 times their raw values to ease comparison with the line drawing case.

Online visual discrimination

This case differs to the others as it is a classification (or decision making) task, where at the end of the presentation of the MNIST image the model must decide which number the digit belongs to (between 0 and 9). Since the decision is only made at the end of the sequence and targets are unavailable at intermediate points, this is a task with hard temporal credit assignment. The output of the model is a vector with probabilities of size 10 (one entry for each number), and the model was trained to maximise the likelihood of the target number using a standard cross-entropy error function.

Visual-language task

The architecture for the caption generation task consists of a pretrained convolutional neural network (CNN) coupled with an RNN (LSTM). The cerebellar network only communicates with the LSTM. The LSTM network has one layer of 256 LSTM units and the cerebellar network has two hidden layers (i.e. here we explicitely model a layer of Granule Cells and one of Purkinje Cells) of 1024 neurons.

The process from image to model-generated caption follows previous work[83] and is described next. As part of image preprocessing and data augmentation, which helps prevent model overfitting, a given image is randomly cropped to size 224×224, flipped horizontally with even chance, and appropriately normalised to be given to a pretrained Resnet model[84]. A feature vector X of size 256 is thus obtained and passed to the LSTM at timestep 0. The LSTM is subsequently presented the "gold standard" caption $\{w_i\}_{i=1}^{n}$ one word per timestep, each time learning to predict the next word (unsupervised task); i.e., at timestep t the model learns $P(w_t|X, \{w_i\}_{i=1}^{t-1})$. The network simultaneously learns a word embedding so that each word w_i is first transformed to a feature vector of size 256 before being given as input to the LSTM (as illustrated in (Fig. 8a). With a preset vocabulary of 9956 distinct words, the final output of the model ($P(w_i)$) is a probability vector of size 9956.

We found the models to be generally prone to overfitting the training data. For this reason, we apply dropout (during training as described by Srivastava et al.[85]) on the input to the LSTM, where a given input element is set to zero with $p = 0.5$ probability. Once training is complete the models can generate their own captions to previously unseen images (Figs. 8, S10). Given an image at timestep 0, the model output at timestep i is the word with the highest probability, and the same word is then provided as input to the model at timestep $i+1$. In this way the model can autonomously output an entire sequence of words which forms a predicted caption. In the (highly) rare case where the model generates a sequence of > 20 words, we consider only the first 20 words as its caption.

We used the COCO training data set ILSVRC-2012-CLS [1][50] which holds 414113 image-caption pairs with 82783

[1] This is a commonly used dataset available for our purposes under a Creative Commons license.

unique images while the held-out validation set (used for Fig. 8b, c) holds 202654 with 40504 unique images; note that each image therefore has $\sim$ 5 distinct gold standard captions. Training takes place in batches of 100 image-caption pairs, with a learning rate of 0.001. Model performance is averaged over 10 random seeds. The performance is quantified in bits per word, which measures how good the model is at predicting the validation set. More specifically if a model assigns high probability to the test set (low BPW) it means it is not surprised by it hence indicating a good understanding of the language.

In order to judge the models beyond their learning curves in BPW, we quantify their ability to generate captions using variety of language modelling metrics popular in the field of language evaluation. In particular, we compare model-generated captions against the gold standard captions using standard metrics in language modelling. We use the Semantic Propositional Image Caption Evaluation or SPICE metric, referred to as caption score. This metric has been shown to be more accurate as it better captures the semantic structure of the generated captions [52].

Our code implementation is based on https://github.com/yunjey/pytorch-tutorial/tree/master/tutorials/03-advanced/image_captioning.

Demixed principal component analysis

To study the response dynamics specific to task variables we perform demixed principal component analysis (dPCA) [86]. Demixed PCA extracts low-dimensional components that explain maximum population variance constrained by task-specific variables, such as the input stimulus. As a result we obtain principal components that are specific to task variables. The simulated neural data we provide as input to dPCA is a three-dimensional array (n, s, t) with neuronal activity (concatenated across seeds), stimulus identity and time, respectively.

Statistical analysis

Because the initial conditions of these type of models influence its learning trajectory we run our models across 10 different randomly chosen seeds. Significance was then tested using a paired t-test across the different seeds on the relative changes (e.g. ccRNN relative to cRNN). Significance levels are represented as * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$) and **** ($p < 0.0001$).

Measuring cerebro-cerebellar feedback similarity

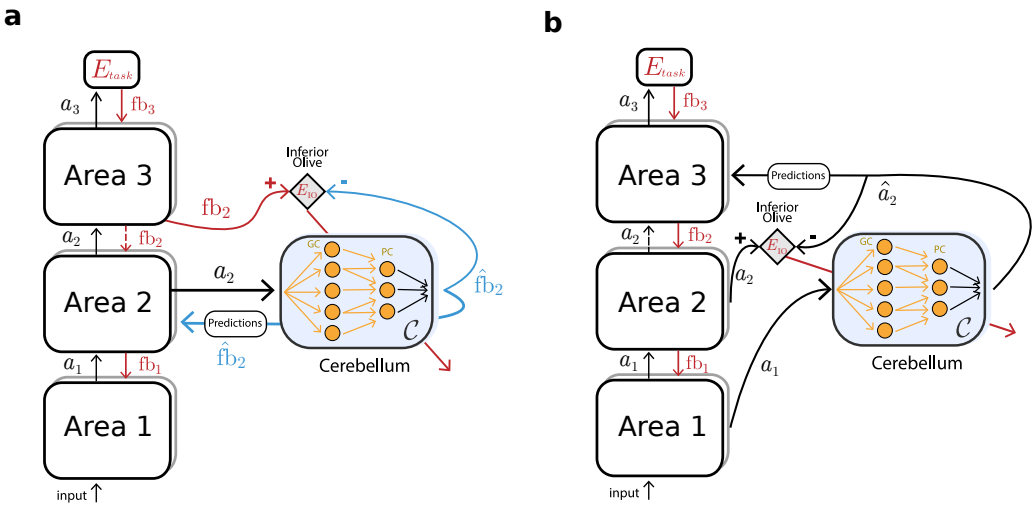
The learning curves of ccRNN plotted against cRNN with a limited feedback horizon highlights the benefit of the feedback predicted by the cerebellar network. This indicates that the predicted feedback can indeed approximating the desired cerebral feedback. To verify this, we quantified the cerebro-cerebellar feedback similarity using cosine similarity - "cossimilarity" - between the predicted feedback and the optimal temporal cerebral feedback (as derived by gradient descent). Specifically given two arbitrary vectors $\mathbf{x}$ and $\mathbf{y}$

$$\text{cossimilarity}(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x} \cdot \mathbf{y}}{||\mathbf{x}||_2 ||\mathbf{y}||_2} \tag{3}$$

where $\mathbf{x}$ is the predicted feedback and $\mathbf{y}$ the true optimal feedback, $\cdot$ denotes the dot product and $||||_2$ the Euclidean norm.

It is important to emphasise that the true feedback is never actually provided to the model (as it goes beyond the feedback horizon K considered). Instead the cerebellum only learns through a combination of cerebral feedback within horizon K and a bootstrapped term (see details above). This measure allow us to evaluate how much information about this ideal feedback can the cerebellum approximate. The final result is shown in Fig. 5a. To provide the reader with intuition about how having external feedback available just at the end, which would lead to a gradual loss of the ability of the cerebellum to make good predictions for earlier points in the task, we highlight two task variants in which the task error is only defined at the end visual discrimination and a simple line drawing variant where the external task feedback is only provided at the end of the task.

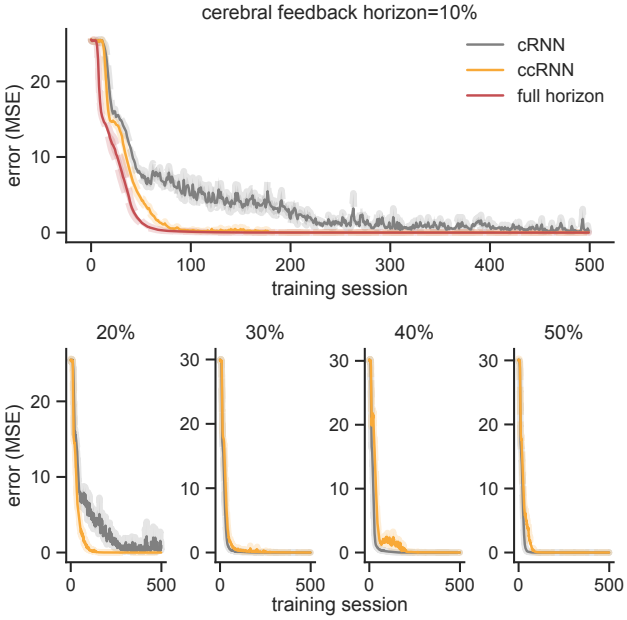
Supplementary Information



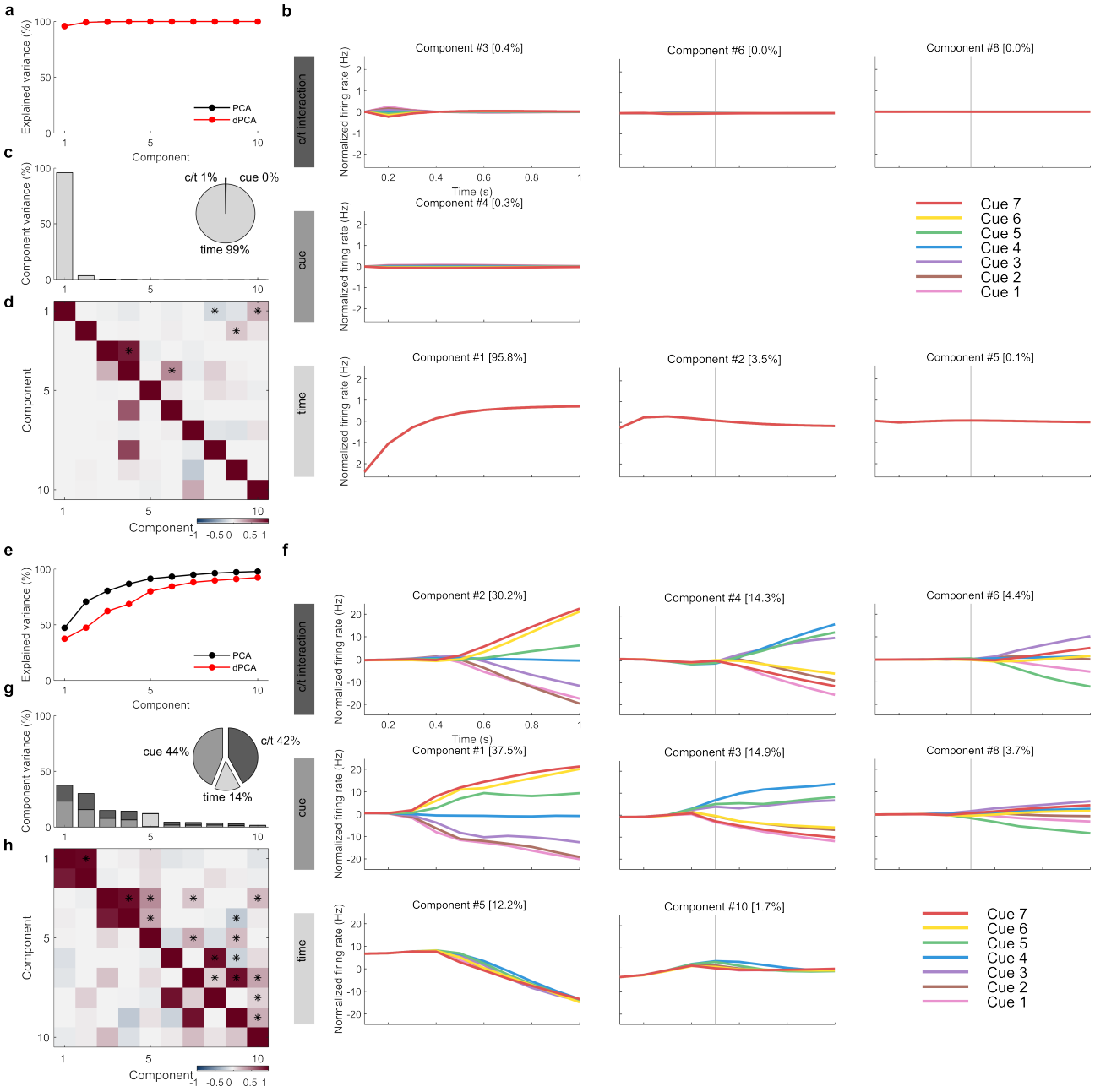
Supplementary Figure S1. Cerebellum as decoupling machine in feedforward multi-area networks. (**a**) Illustration of decoupling feedback processing. The cerebellum makes predictions of the feedback expected by brain area 2, decoupling the main network from downstream brain areas (dashed red arrow). (**b**) Case of decoupling feedforward processing. The cerebellum predicts the forward activity expected by brain area 3, thereby approximating (and decoupling) the forward computations between brain area 1 and 3 (dashed black arrow). Note that the cerebellum could, in principle, approximate feedback and feedforward processing across many more brain areas (i.e. brain area 2 could be expanded in multiple brain areas).

Supplementary Table S1. Relationship between the internal models of the cerebellum with decoupling machines [33]. The properties of the forward model of the cerebellum can be set against those of feedback decoupling (blue); similarly, the properties of the inverse model of the cerebellum can be set against those of forward decoupling (red). The internal models here focus on the classical motor control setting but can be extended to cognition, where for example a "mental model" replaces the "controlled object" [24]. Abbreviations: MM, main model; temp., temporal; spat. spatial.

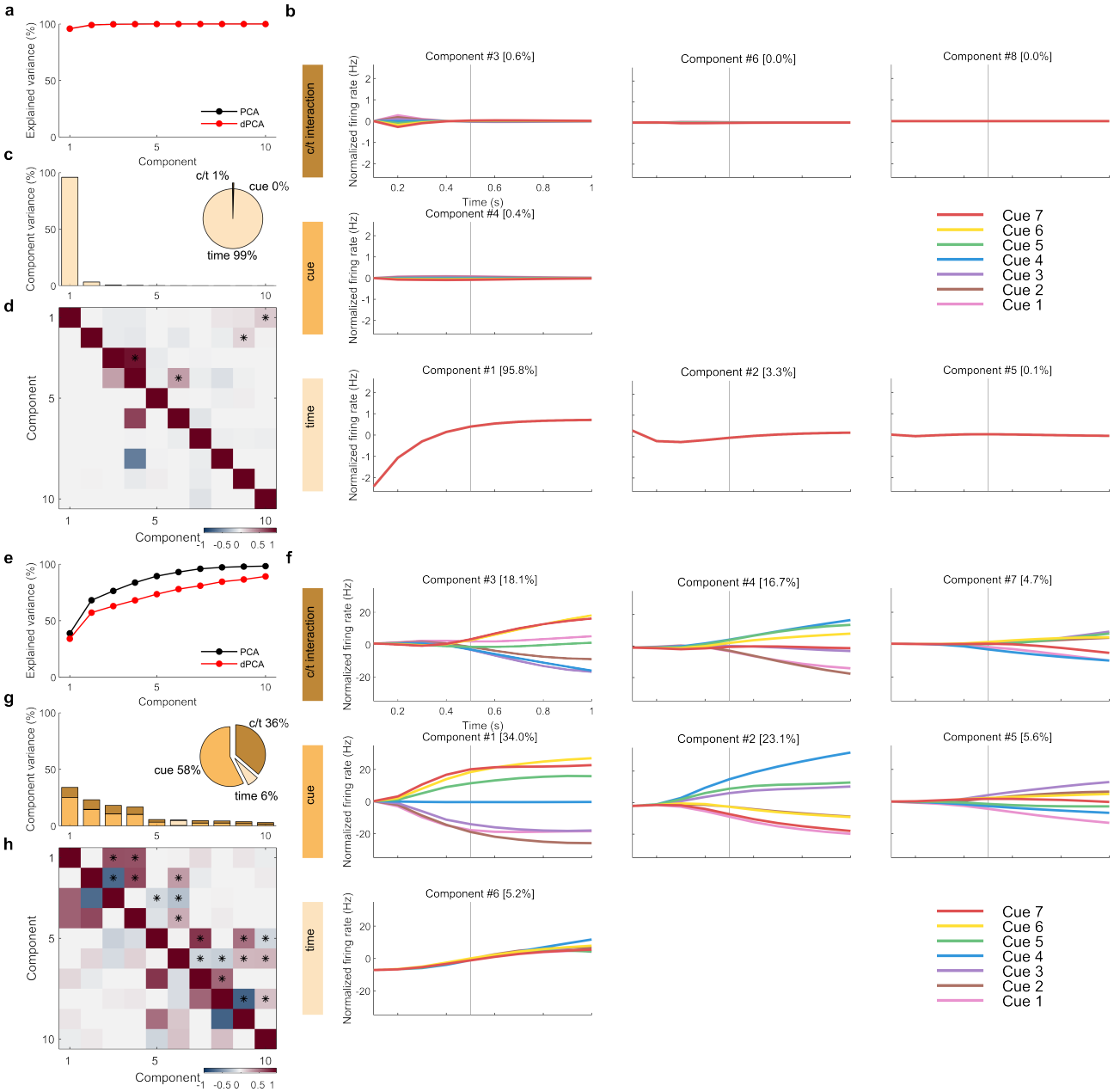
	Forward Model	Feedback Decoupling	Inverse Model	Forward Decoupling
controller	cerebral (motor) cortex	main model (MM)	cerebellum	synthesiser
input	motor state/command	area state*	sensory/desired state	(temp.) area state* (spat.) upstream state*
output prediction	future state	(temp.) future gradient (spat.) downstream gradient	motor command	(temp.) future state (spat.) downstream state
output destination	cerebral (motor) cortex	MM: same area	controlled object	(temp.) MM: same area (spat.) MM: downstream area



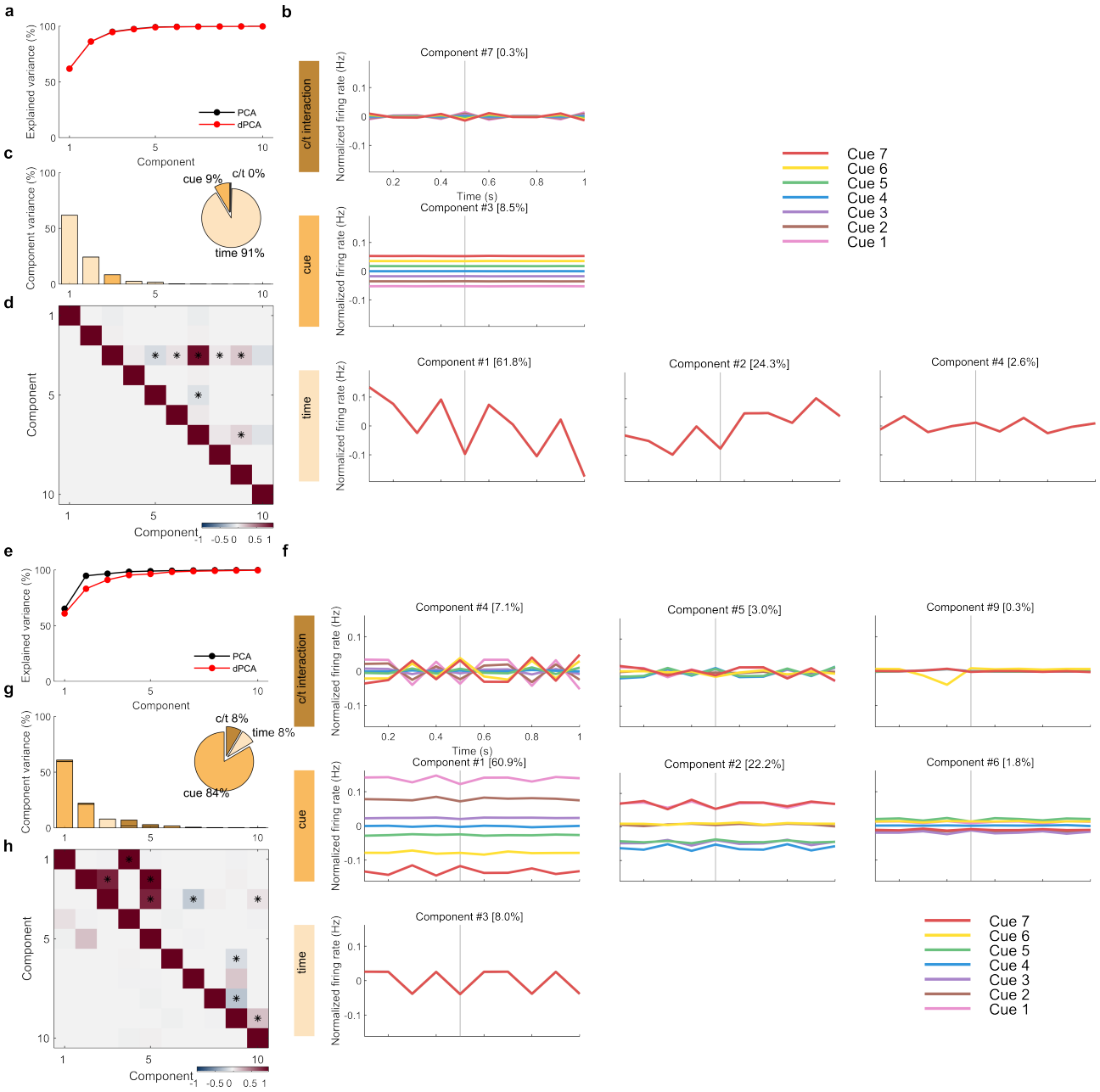
Supplementary Figure S2. Learning for different cerebral feedback horizons for the line drawing task (cf. Fig. 2d). Feedback horizon is given as percentage of task duration (10 time steps). Results presented in main text (Fig. 2b) shown on top row along with RNN trained with full horizon (i.e. cerebral feedback horizon = 100%).



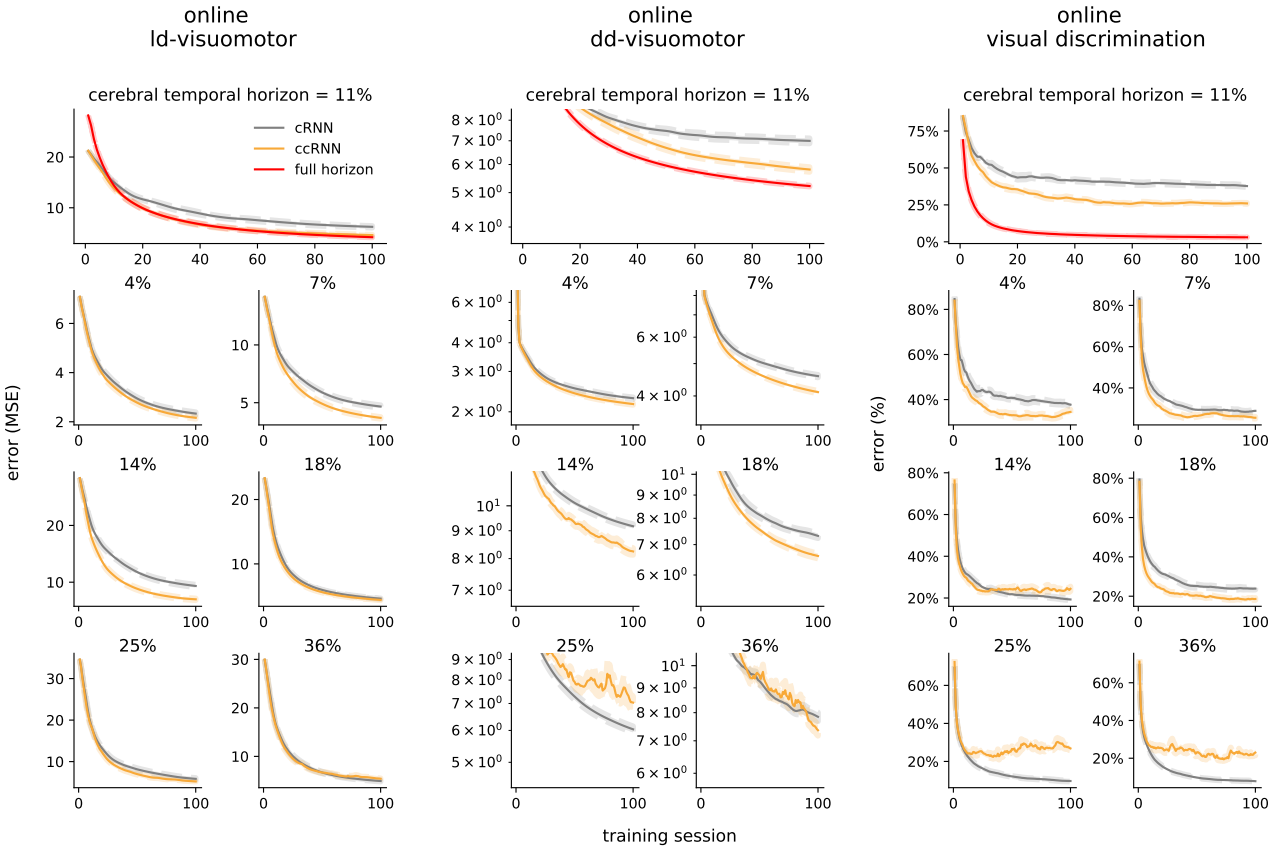
Supplementary Figure S3. Demixed PCA of cRNN network at the beginning and end of learning (cf. Fig. 2e,f). Early and late learning corresponds to training session 1 (top a-d) and 200 (bottom e-h), respectively. (**a, e**) Cumulative variance explained by PCA (black) and dPCA (red) components. (**b, f**) Demixed principal components for cue, time and cue/time interaction task variables. In each subplot there are 7 lines corresponding to the 7 cues (cf. Fig. 2a). (**c, g**) Explained variance for individual demixed principal components. Pie chart shows how the total variance is split between different task variables. (**d, h**) Dot product between all pairs of the first 15 demixed principal components (upper-right triangle) and correlations between all pairs of the first 10 demixed principal components (bottom-left triangle). Stars denote statistical significance ($p < 0.05$).



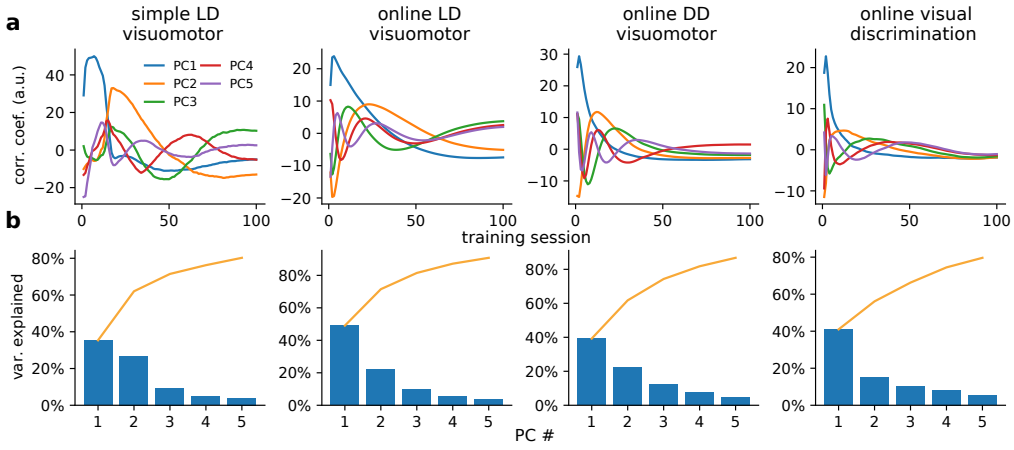
Supplementary Figure S4. Demixed PCA of ccRNN cerebral network at the beginning and end of learning (cf. Fig. 2e,f). Early and late learning corresponds to training session 1 (top a-d) and 200 (bottom e-h), respectively. (**a, e**) Cumulative variance explained by PCA (black) and dPCA (red) components. (**b, f**) Demixed principal components for cue, time and cue/time interaction task variables. In each subplot there are 7 lines corresponding to the 7 cues (cf. Fig. 2a). (**c, g**) Explained variance for individual demixed principal components. Pie chart shows how the total variance is split between different task variables. (**d, h**) Dot product between all pairs of the first 15 demixed principal components (upper-right triangle) and correlations between all pairs of the first 10 demixed principal components (bottom-left triangle). Stars denote statistical significance ($p < 0.05$).



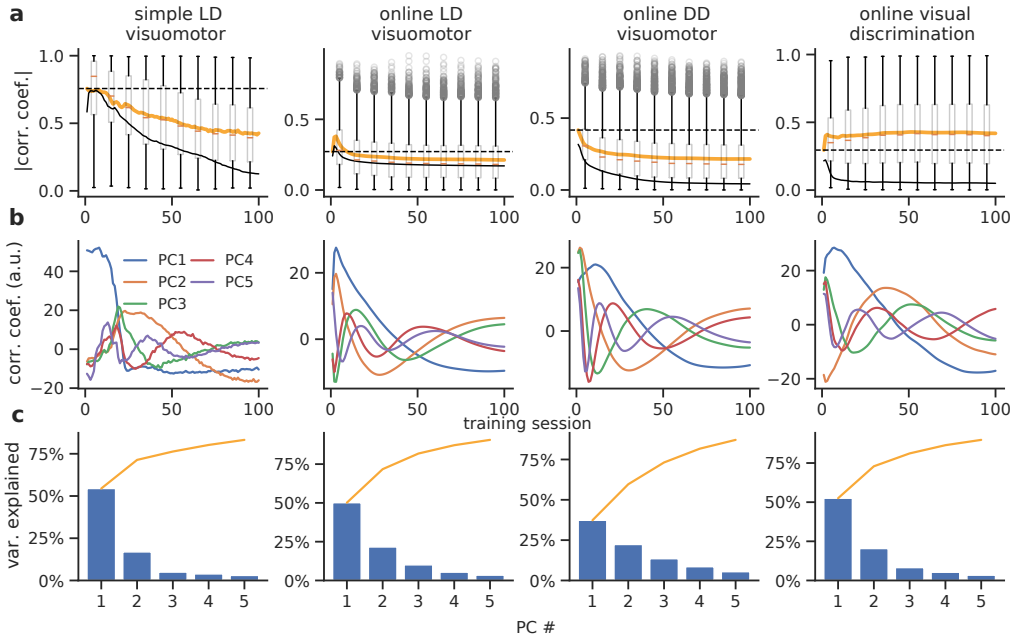
Supplementary Figure S5. Demixed PCA of ccRNN cerebellar network at the beginning and end of learning (cf. Fig. 2g,h). Early and late learning corresponds to training session 1 (top a-d) and 200 (bottom e-h), respectively. (**a, e**) Cumulative variance explained by PCA (black) and dPCA (red) components. (**b, f**) Demixed principal components for cue, time and cue/time interaction task variables. In each subplot there are 7 lines corresponding to the 7 cues (cf. Fig. 2a). (**c, g**) Explained variance for individual demixed principal components. Pie chart shows how the total variance is split between different task variables. (**d, h**) Dot product between all pairs of the first 15 demixed principal components (upper-right triangle) and correlations between all pairs of the first 10 demixed principal components (bottom-left triangle). Stars denote statistical significance ($p < 0.05$).



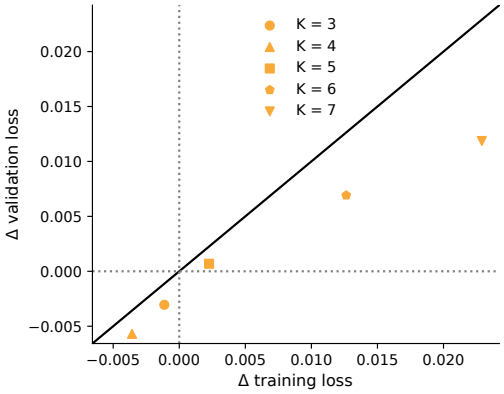
Supplementary Figure S6. Learning for different cerebral feedback horizons for the online visuomotor and discrimination tasks (cf Fig. 3d). Feedback horizon is given as percentage of task duration (28 timesteps). Results presented in main text (Fig. 3b) shown on top row along with RNN trained with full horizon (i.e. cerebral feedback horizon = 100%).



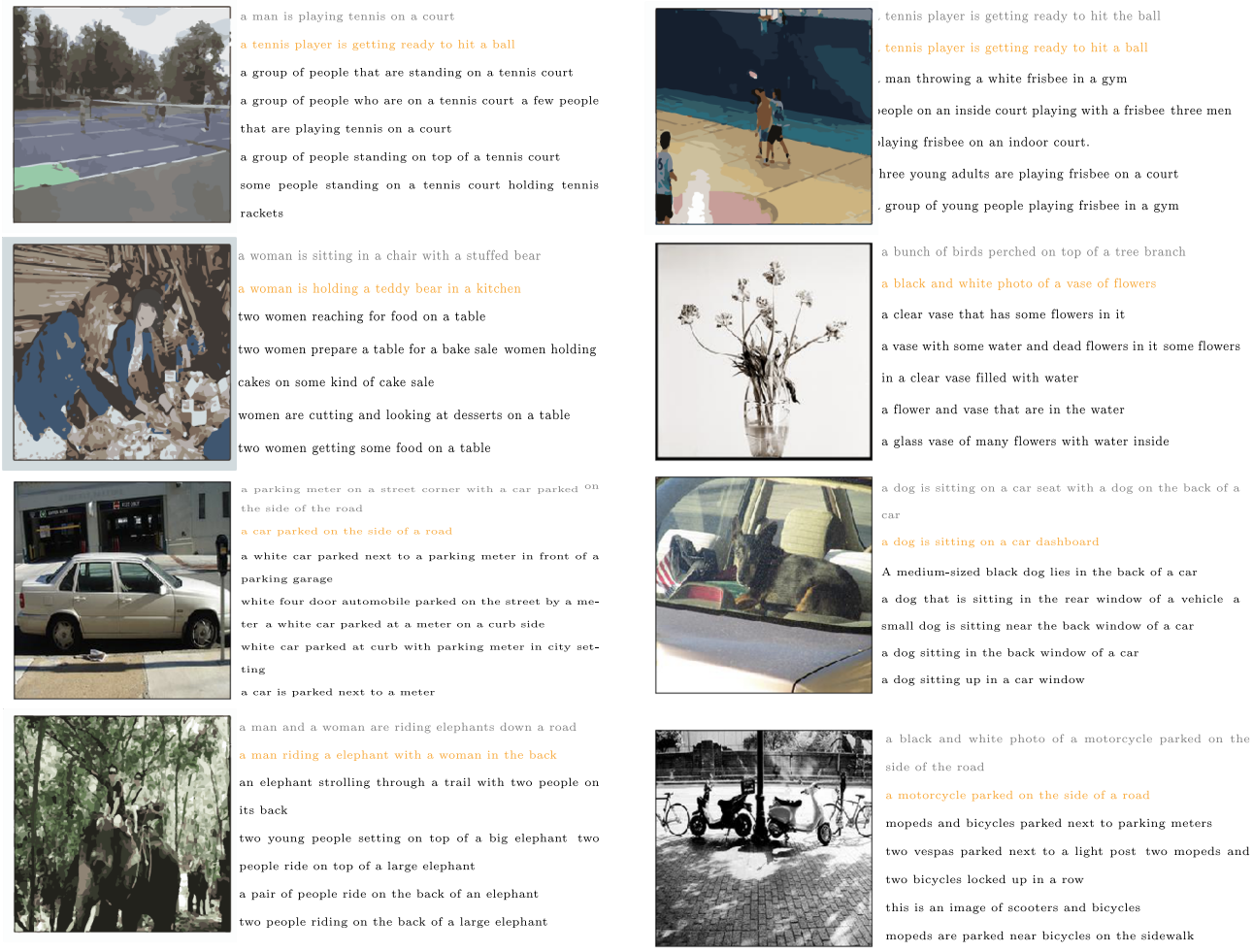
Supplementary Figure S7. Pair-wise correlations over learning. (**a**) extension of Fig. 6b for top 5 principal components. (**b**) Variance explained by each component (accumulation in orange).



Supplementary Figure S8. Pair-wise correlations over learning with a *fixed* cerebellar module. (**a**) Box plot showing the mean and distribution of pair-wise cerebro-cerebellar correlations over learning. Mean correlation coefficient for the fully plastic ccRNN model (solid black line) and fully fixed ccRNN (i.e. without any form of plasticity in both cerebral and cerebellar networks; dashed black line) are given for reference. (**b**) Top 5 principal components. (**c**) Variance explained by each component (accumulation in orange).



Supplementary Figure S9. Generalisation of ccRNN (orange scatter) for feedback horizons *K* from 3 to 7. The change in loss is computed with reference to the cRNN (i.e. ccRNN - cRNN). Training loss is calculated *after* training for a fair comparison with final validation performance.



Supplementary Figure S10. Example images and captions from the validation set with corresponding model captions (cRNN in grey and ccRNN in orange) and gold standard captions (black). Here we show a combination of examples of how the models describe the presented image. In some case all or some models fail to give an accurate description of the image. In other cases all models are able to provide an accurate caption for the image, with each model displaying subtle differences in the generated captions.