Remembrance of things practiced: Fast and slow learning in cortical and subcortical pathways

James M. Murray and G. Sean Escola
Zuckerman Mind, Brain, and Behavior Institute, Columbia University

January 22, 2020

The learning of motor skills unfolds over multiple timescales, with rapid initial gains in performance followed by a longer period in which the behavior becomes more refined, habitual, and automatized. While recent lesion and inactivation experiments have provided hints about how various brain areas might contribute to such learning, their precise roles and the neural mechanisms underlying them are not well understood. In this work, we propose neural- and circuit-level mechanisms by which motor cortex, thalamus, and striatum support such learning. In this model, the combination of fast cortical learning and slow subcortical learning gives rise to a covert learning process through which control of behavior is gradually transferred from cortical to subcortical circuits, while protecting learned behaviors that are practiced repeatedly against overwriting by future learning. Together, these results point to a new computational role for thalamus in motor learning, and, more broadly, provide a framework for understanding the neural basis of habit formation and the automatization of behavior through practice.

Introduction

The acquisition and retention of motor skills, for example shooting a basketball or reaching for a piece of food, are crucial for humans and other animals. Such skills are acquired through repeated practice, and the effects of such practice are numerous and complex [1, 2, 3]. Most obviously, practice leads to improved performance, e.g. as measured by movement accuracy [4]. In addition, practice leads to the formation of habits, which can be defined operationally as decreased sensitivity of a behavior to goals and rewards [5, 6]. Finally, practice leads to behavioral automatization, which can be measured by improved reaction times or decreased cognitive effort [2]. In many cases, these multiple effects of practice take place on different timescales, with large gains in performance obtained early in learning, followed by a much longer period in which the learned behavior becomes more refined, habitual, and automatic [7, 8]. This suggests the possibility that motor learning can be divided into a fast, overt learning process leading to improved performance, combined with a slower, covert learning process with subtler effects related to habit formation and behavioral automatization. It is currently not well understood, however, whether these multiple learning processes that occur during practice are truly distinct or merely different facets of the same underlying process, whether they operate sequentially or in parallel, and how they might be implemented in the brain.

Neurobiologically, the descending pathway through the sensorimotor region of the striatum, the major input nucleus of the basal ganglia, plays an important role in both the learning and execution of learned motor behaviors [9, 10, 11], as well as in habit formation for highly practiced behaviors [12, 13]. Such learning is facilitated by dopamine-dependent plasticity at input synapses to striatum, where this plasticity plays a

role in reinforcing behaviors that have previously led to reward [14]. The sensorimotor striatum is able to influence behavior through two pathways: one descending pathway to motor-related brainstem nuclei [15] and another pathway to motor cortex via the basal ganglia output nuclei and motor thalamus [16]. Neural activity in sensorimotor striatum itself is driven by input from both cortical and subcortical sources. The dominant cortical input comes from sensorimotor cortex [17, 18], and a great deal of research has implicated this pathway in motor learning [19, 20]. The dominant subcortical input, meanwhile, comes from thalamus, including the rostral intralaminar nuclei and parafascicular nucleus [21, 22]. Although this subcortical input to striatum has received less attention than the corticostriatal pathway, it has also been shown to represent behaviorally relevant variables [23, 24, 25] and to be important for the production of learned behaviors [26]. In general, however, the role of this thalamic pathway in motor learning and how it might differ from the corticostriatal pathway is not well understood.

In order to obtain some understanding about the roles that cortex, thalamus, and striatum play in motor learning, experimental manipulations of these structures and of the projections between them during different stages of learning provide useful constraints. Recent lesion and inactivation experiments in rats have shown that sensorimotor striatum and its thalamic inputs are necessary for both learning and execution of a skilled lever-press sequence [27]. In the same task, however, input from motor cortex to striatum is necessary for learning the task, but not for performing the task once it has already been learned [28, 27]. Furthermore, the representation of task-related kinematic variables in striatum during the task is similar in rats with and without cortical lesions [29]. Evidence for the gradual disengagement of motor cortex during the learning of skilled movements has also been found in mice [30] and humans [31, 32]. Together, the above results provide hints about the roles played by cortex, thalamus, and striatum in the acquisition and execution of skilled movements, suggesting that responsibility for driving learned behaviors may be gradually offloaded from cortical to subcortical circuits through practice. It is not immediately obvious, however, which neural- or circuit-level mechanisms might explain these results, nor how they might relate to the multiple effects of practice discussed above.

In this work, we propose a mechanistic theory of learning at cortico- and thalamostriatal synapses that provides a unified description of the lesion and inactivation results described above and, more broadly, points to computationally distinct roles for the two inputs to striatum in the acquisition and retention of motor skills. As a starting point for addressing learning, retention, and recall accuracy, we begin by modeling a single neuron within striatum. We mathematically derive the neuron's forgetting curve, which quantifies the probability that the neuron produces the correct output in response to a particular pattern of inputs as a function of how far in the past learning occurred. If the neuron's input synapses are all modified by a supervised (i.e. error-minimizing) plasticity rule, then the neuron learns to produce the correct output in response to a particular input, but there is no added benefit to repeated practice once the output has been learned. If, on the other hand, some of the neuron's input synapses are modified by a slower, associative learning rule, then patterns that are practiced many times become far more resistant to being overwritten by future learning. We interpret the first input, with fast supervised learning, as motor cortex, and the second input, with slow associative learning, as thalamus. We further show that, as a behavior is practiced, there is a gradual and covert transfer of control, with the activity of the downstream population increasingly driven by the second input pathway. In addition to providing a unified description of the lesion and inactivation experiments described above, this model generates predictions for the expected effects of perturbations to the cortical and thalamic pathways. Further, it proposes a new computational role for thalamic inputs to striatum and, more broadly, provides a framework for understanding habit formation and the automatization of behavior as the control of behavior is transferred from cortical to subcortical circuits.

Results

In the following sections, our aim will be to develop a theory that describes the effects of learning and practiced repetition described above. As a first step, we will mathematically investigate learning and forgetting within a simple model of supervised learning in a single striatal neuron receiving cortical inputs. This will provide a minimal quantitative theory for addressing questions about retention and accuracy in learned behaviors, but will not yet account for effects related to practiced repetition. In order to account for such effects as well as for the lesion and inactivation results described above, we will next add a second set of inputs from thalamus. By endowing these inputs with associative learning, we will show that this second input pathway leads to enhanced retention and greater accuracy for patterns that are practiced many times during training. These two types of learning together give rise to a covert learning process in which control of the striatal population is gradually transferred from the cortical to the thalamic inputs through repeated practice. Finally, we describe how this two-pathway model leads to a number of experimental predictions about the expected effects of perturbations and inactivations of the cortical and thalamic inputs.

A single-neuron model describes learning and forgetting

In order to develop a quantitative theory of learning at the most basic level, we began by studying a classical model of a single neuron: the perceptron [33], which in our model corresponds to a single striatal neuron receiving cortical input (Figure 1a,b). In this model, N_x input signals x_i are multiplied by synaptic weights w_i and then summed. If the summed value is greater than 0, the neuron's output is +1, otherwise the output is -1. This is described by the equation $z^\mu = \mathrm{sgn}(\mathbf{w}^\mu \cdot \mathbf{x}^\mu)$, where μ denotes one of P possible input patterns (Figure 1b). The goal is then for the synaptic weights $\mathbf{w}$ to be adjusted so that, for each pattern μ, the output matches a target value $\hat{z}^\mu$. Thus, the task is binary classification, in which random input patterns ($x_i^\mu \sim \mathcal{N}(0,1)$, where $\mathcal{N}$ denotes the standard normal distribution) are mapped onto random output values ($\hat{z}^\mu = \pm 1$). In our neurobiological interpretation, the perceptron corresponds to a single neuron in striatum, the major input structure of the basal ganglia. This neuron receives inputs from cortex with information about cues, context, and sensory feedback, and it learns to adjust the strengths of its input weights so that the neuron is active in response to some inputs and inactive in response to others.

This learning problem can be illustrated geometrically in the N_x-dimensional space of synaptic weights, where the weight vector $\mathbf{w}^\mu$ just before learning pattern μ is represented as a vector in this space (Figure 1c). In order for the classification to be correct, the weight vector after training should have positive overlap with the vector $\hat{z}^\mu \mathbf{x}^\mu$, which defines a classification boundary (red line in Figure 1c), with all weight vectors on one side of this boundary leading to correct classification. This can be effected by updating the weights to $\mathbf{w}^\mu + \Delta\mathbf{w}^\mu$, where

$$\Delta\mathbf{w}^\mu = \begin{cases} (\kappa\hat{z}^\mu - u^\mu)\mathbf{x}^\mu / N_x, & u^\mu \hat{z}^\mu < \kappa, \\ 0, & \text{else}, \end{cases} \tag{1}$$

where $u^\mu = \mathbf{w}^\mu \cdot \mathbf{x}^\mu$ is the summed input to the neuron [34]. The first line of this equation says that, if the classification is initially incorrect or correct but within a margin κ of the classification boundary (dashed line in Figure 1b), then an update is made such that the new weight vector $\mathbf{w}^\mu + \Delta\mathbf{w}^\mu$ lies on the correct side of the classification boundary (Figure 1c) with margin κ. The second line, on the other hand, says that, if the classification is initially correct with a margin of at least κ, then no update needs to be made. Because κ can be absorbed into an overall rescaling of $\mathbf{w}$, we henceforth set $\kappa = 1$.

Clearly a potential problem exists whenever more than one pattern is being learned, however. After applying (1) to ensure correct classification of pattern μ, followed by applying the same update rule for

pattern $\mu+1$, there is no guarantee that the weight vector will still be on the correct side of the classification boundary for pattern μ. That is, the learning of new information in this model has the potential to interfere with and overwrite previously learned information. In the standard paradigm for learning in the perceptron, this problem is solved by cycling many times through all of the P patterns to be learned. In one of the most celebrated results in computational neuroscience, it has been shown that, in the large-N limit, a weight vector producing correct classification of all patterns can be obtained whenever $P < 2N_x$ [35, 36].

In order to address effects related to retention and recall, we instead study the case in which the patterns are learned in sequence, so that no further training with pattern μ occurs after training with pattern $\mu+1$ has begun. Concretely, we first train the perceptron by applying the update rule (1) for each of the P patterns in the sequence. We then test the perceptron by testing the classification of each pattern using the final weight vector $\mathbf{w}^P$. Intuitively, we expect that the most recently learned classifications will remain correct with high probability during testing since the weight vector will have been updated only a small number of times after learning these patterns. On the other hand, for classifications learned much earlier and hence overwritten by a large number of subsequent patterns, the probability of incorrect classification should approach chance level.

In order to describe this process of training, overwriting, and testing quantitatively, we calculated the probability of incorrect classification (i.e. the error rate) for each pattern during testing after all patterns have been trained. The full calculation is presented in Appendix 1. To summarize, it begins by assuming the weight vector is $\mathbf{w}^\nu$ just before learning some generic pattern ν, then applies the update rule (1) using randomly drawn $\mathbf{x}^\nu$ and $\hat{z}^\nu$ to obtain $\Delta\mathbf{w}^\nu$. For the following $P-\nu$ steps ($\mu = \nu+1, \ldots, P$), the updates $\Delta\mathbf{w}^\mu$ will be in directions that are random with respect to $\mathbf{w}^\nu$. Hence, the evolution of the weight vector from $\mathbf{w}^\nu$ to $\mathbf{w}^P$ can be described as a drift-diffusion process characterized by the probability distribution $p(\mathbf{w}^P|\mathbf{w}^\nu + \Delta\mathbf{w}^\nu)$. Once the probability distribution for $\mathbf{w}^P$ is known, the error rate during testing is calculated by averaging $z^\nu = \mathrm{sgn}(\mathbf{w}^P \cdot \mathbf{x}^\nu)$ over $\mathbf{w}^P$, as well as over the initial weight vector $\mathbf{w}^\nu$ and the random input pattern $\mathbf{x}^\nu$.

The result of this calculation is that the error rate during testing of pattern ν is given by

$$p(z^\nu \neq \hat{z}^\nu) = F(\gamma), \tag{2}$$

where the form of the function F is given in Appendix 1, and $\gamma = e^{-q(P-\nu)/N_x}$, where $q \simeq 0.798$ is the probability that a nonzero update is made according to (1) during any particular step of training. This result defines the single-neuron forgetting curve, which is shown in Figure 1d. (While the pattern index ν is the independent variable in the forgetting curve, we plot the curve as a function of $(P-\nu)/N_x$ so that earlier times appear on the left and later times on the right, and so that the curve is independent of P, which may be arbitrarily large.) As expected, the forgetting curve in Figure 1d shows that recently learned classifications are recalled with very low error rates, while those learned far in the past are recalled with error rates approaching chance level. We can further observe that, because the number of synapses N_x appears only through the parameter γ, the number of patterns that can be learned with error rate at or below any particular threshold is proportional to N_x. This means that the memory is extensive in the number of synapses, as in the classical result described above for the perceptron with cycled training. Finally, because the perceptron implements learning in a single step and may therefore be questionable as a neurobiological model of learning, we show using simulations that a similar forgetting curve is obtained using a gradient-descent learning rule, in which small updates are accumulated over many repetitions to minimize the readout error (Supplemental Figure 1).

The single-neuron model that we have studied in this section provides a zero-parameter baseline theory for addressing questions related to learning, retention, and recall accuracy. The ability of this model to

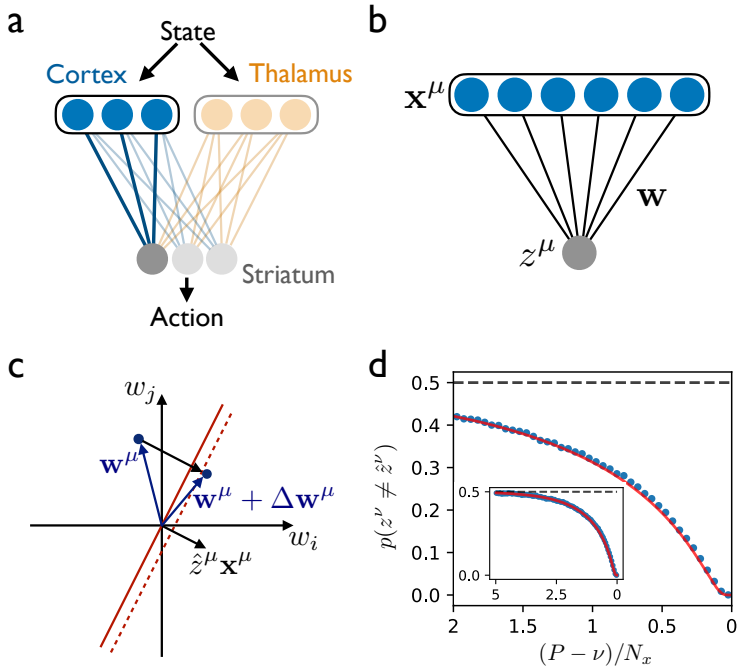


Figure 1: The forgetting curve for a single neuron with supervised learning. **(a)** A sensorimotor circuit model in which cortex and thalamus contain state information (e.g. cues and sensory feedback) and provide input to striatum, which learns to drive actions that lead to reward. **(b)** A model of a single striatal neuron receiving cortical inputs, where weights $\mathbf{w}$ are trained such that random input patterns $\mathbf{x}^\mu$ produce the correct classifications $z^\mu = \hat{z}^\mu$. **(c)** The input vector $\mathbf{x}^\mu$ defines a hyperplane in the space of weights $\mathbf{w}$, and the update rule modifies the weight vector to give the correct classification with margin $\kappa = 1$. **(d)** The probability of incorrect classification when testing pattern ν after learning $P = 2N_x$ patterns sequentially. The most recently learned patterns are on the right, with earlier patterns on the left. The solid curve is the theoretical result; points are simulated results. Inset shows the same result for a perceptron trained with $P = 5N_x$ patterns.

learn new information by overwriting recently learned information is in contrast to the case of the classical perceptron, in which patterns are cycled through repeatedly during training, leading to a classification error rate of zero for up to $P = 2N_x$ patterns. However, no further classifications could be learned in this case without drastically impairing the classification performance of the already-learned patterns. This phenomenon, known as catastrophic forgetting [37], is obviated in the sequentially trained model studied here. As previously pointed out in theoretical studies of related models [38, 39], although sequentially trained neurons and neural networks have a smaller overall memory capacity than models in which training is repeated, they have the advantage that learned patterns decay smoothly over time as new patterns are learned, even as $P/N_x \to \infty$.

A two-pathway model describes the effects of practice

In the preceding section, we derived the forgetting curve for a single-neuron model of a striatal neuron receiving cortical inputs with supervised learning in order to quantitatively address learning and forgetting. Because learning of any given pattern occurs in a single step and stops once the error is zero, however, this simple model is unable to describe the effects of practice. In order to study the effects of practice, we made the simplest possible addition to the model by including a second input pathway, where synaptic weights in this pathway are modified with an associative update rule, as originally proposed by Hebb [40] (Figure 2a). Specifically, these synapses are strengthened by coactivation of their inputs and outputs, so that the same inputs will in the future tend to produce the same outputs. This strengthening, which is independent of errors or rewards, continues to accumulate with each repetition if an input pattern is presented many times. Importantly, the Hebbian learning should be slow compared to the supervised learning, so that it will not strengthen an incorrect input-output association before the supervised learning has taken place. Intuitively, then, this picture suggests that repeated training of a particular classification should strengthen the association and make it more resistant to being overwritten by later learning. We shall show below that this is indeed the case. After developing the theory for this two-pathway model, we shall show in the following section that identifying the second input to striatum as thalamus provides an explanation for the asymmetric effects of motor cortical and thalamic lesions in rats described in the Introduction [28, 29, 27].

Concretely, we extended the single-neuron model by adding a second set of inputs $\mathbf{y}^\mu$, where the N_y inputs $y_i^\mu \sim \mathcal{N}(0,1)$ are random and independent of the first set of N_x inputs x_i^μ. The output is then given by $z^\mu = \mathrm{sgn}(\mathbf{w}\cdot\mathbf{x}^\mu + \mathbf{v}\cdot\mathbf{y}^\mu)$, where the synaptic weights $\mathbf{v}$ are updated according to a Hebbian learning rule. Specifically, if $\mathbf{v}^\mu$ is the weight vector just before training on pattern μ, then the weight vector just after training on pattern μ is given by $\mathbf{v}^\mu + \Delta\mathbf{v}^\mu$, where

$$\Delta\mathbf{v}^\mu = -\frac{\alpha n_\mu}{N_y \bar{n}}\mathbf{v}^\mu + \sqrt{2}\frac{\beta n_\mu}{N_y \bar{n}}\hat{z}^\mu \mathbf{y}^\mu. \tag{3}$$

The second term in this equation is a modification to the synaptic weight that is proportional to the product of the pre- and post-synaptic activity, making this a Hebbian learning rule. Because the learning of $\mathbf{w}$ is assumed to be fast relative to this Hebbian learning, the downstream activity will have attained its target value by the time the Hebbian learning has any significant effect, allowing us to use $\hat{z}^\mu$ rather than z^μ in (3). The n_μ appearing in (3) is interpreted as the number of times that the behavior μ is repeated when it is trained, and $\bar{n}$ is the average of n_μ over all patterns. The first term in (3) causes the weights to decrease slightly in magnitude with each repetition of each pattern, so that they do not become arbitrarily large after a large number of classifications have been learned. The constants α and β control how quickly the weights decay and how strong the initial updates are, respectively. The weights $\mathbf{w}$, meanwhile, are updated as before, except that now the total summed input appearing in (1) is given by $u^\mu = \mathbf{w}\cdot\mathbf{x}^\mu + \mathbf{v}\cdot\mathbf{y}^\mu$. In our

neurobiological interpretation, the second pathway corresponds to thalamic inputs to the striatal neuron. As suggested by experiments [23, 24, 25], these inputs, like cortical inputs, may convey task-relevant information about cues and sensory feedback. Unlike the cortical inputs, however, the thalamic inputs connect to the readout neuron with synapses that are modified by an associative update rule.

For this two-pathway model, we derived a generalized forgetting curve using the same approach as that used for the perceptron, describing the evolution of both weight vectors $\mathbf{w}$ and $\mathbf{v}$ as drift-diffusion processes and computing statistical averages over weights and patterns. The generalized forgetting curve for the two-pathway model has the form

$$p(z^\nu \neq \hat{z}^\nu) = G\left(\gamma, \rho, \alpha, \beta, \frac{n_\nu}{\bar{n}}\right), \tag{4}$$

where $\rho = e^{-\alpha(P-\nu)/N_y} = \gamma^{\alpha N_x/(qN_y)}$, and the function G is derived in Appendix 2. From (4), we see that the forgetting curve now additionally depends on the Hebbian learning rate β, the ratio N_y/N_x, the Hebbian decay rate α, and the number of repetitions n_ν for each pattern.

In Supplemental Figures 2 and 3, we fixed n_ν to be constant for all patterns and investigated the effects of α, β, and N_y/N_x on the forgetting curves. We found that, while adding the second pathway does not in general improve the memory performance per synapse (of which there are $N_x + N_y$), it does improve the memory performance per *supervised* synapse (of which there are N_x). Further, when the Hebbian decay rate α becomes small, the number of very old patterns that can be correctly recalled is improved, even if normalizing by the total number of input synapses (i.e. by $N_x + N_y$, rather than by N_x).

We next investigated the case in which some patterns are repeated more than others during training. In this case, we found that repeating particular patterns consecutively during training causes those patterns to be retained with lower error rates and for much longer than nonrepeated patterns. If trained with enough repetitions, particular patterns can be recalled essentially perfectly even after $\sim (N_x+N_y)$ additional patterns have been learned (Figure 2b). Further, as long as the number of highly practiced patterns is much less than $N_x + N_y$, the error rate for the remaining patterns is not significantly increased (solid line versus dotted line in Figure 2b). This shows that it is possible for a neuron to become an "expert" at a small number of classifications without performing significantly worse on other classifications that have recently been learned.

We next investigated the dependence of the error rate for a particular pattern ν^* on the interval (the "testing interval") between training and testing, as well as on the number of repetitions n_{ν^*}, for that pattern, again assuming that all other patterns are presented only once during training. We found that the effect of practice is to shift the forgetting curves in a roughly parallel manner (Figure 2c). This parallel shifting of forgetting curves via practiced repetition is a well known result from experimental psychology studies in human subjects performing memory tasks [41, 42, 43, 44, 45]. Plotting the data instead as a function of the number of repetitions n_{ν^*} for a fixed testing interval shows the error rate smoothly decreasing with practice (Figure 2d), again bearing similarity to results from memory studies in humans [46]. This shows that the error rate during testing, for a fixed recency of the practiced pattern, decreases as a function of how many times that pattern was practiced. We next asked, if just a single pattern ν^* is repeated multiple times while all other patterns are trained just once, how many times must pattern ν^* be repeated in order to obtain an error rate below a threshold p_θ during later testing of that pattern? The number of necessary repetitions was found to be a supralinear function of the interval between the training and testing of pattern ν^* (Figure 2e). This dependence could potentially be (and, to our knowledge, has not already been) measured experimentally, for example in human subjects learning paired associations (as in, e.g., Refs. [44, 45]), with certain associations presented multiple times in succession during training, and with varying intervals between training on these repeated patterns and testing.

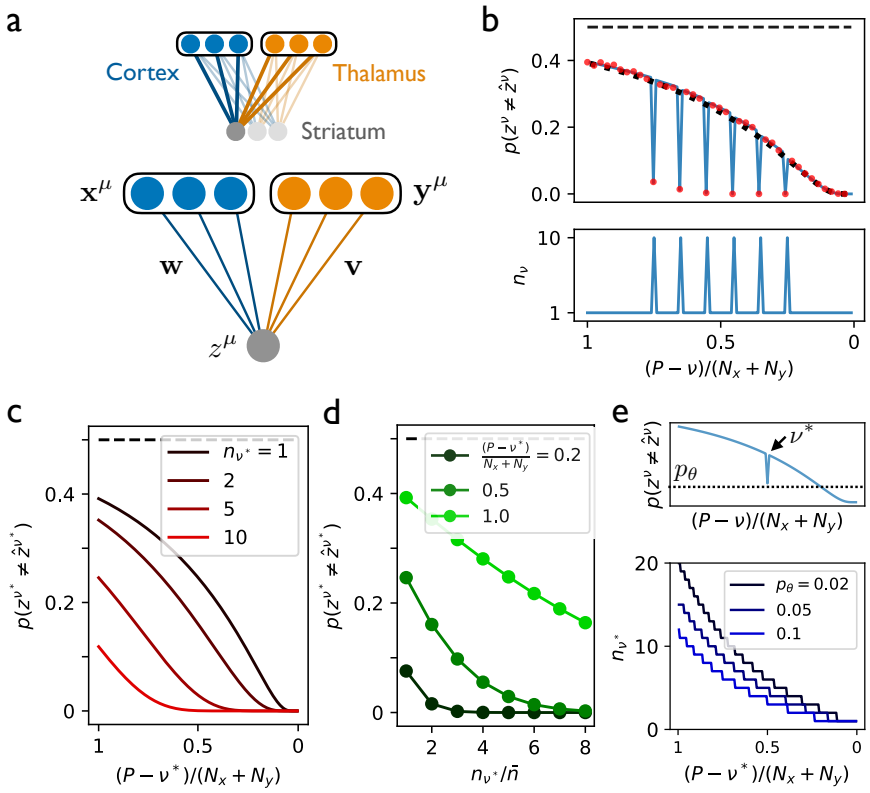


Figure 2: Repetition during training improves recall in the two-pathway model. **(a)** The two-pathway architecture, with cortical and thalamic inputs to striatum (*top*), where fast supervised learning of corticostriatal weights $\mathbf{w}$ is accompanied by slow Hebbian learning of thalamostriatal weights $\mathbf{v}$ (*bottom*). **(b)** The forgetting curve (*top*) in a case where six patterns are repeated multiple times during training, while all other patterns are presented only once (*bottom*). Solid curve is theoretical result ($\alpha = 1, \beta = 1$); points are simulations with $N_x = N_y = 1000$; dotted curve shows the case in which no patterns are repeated multiple times. **(c)** The error rate for pattern ν^*, which is repeated n_{ν^*} times during training, while other patterns are presented only once. **(d)** The error rate as a function of n_{ν^*}, with curves corresponding to different choices of ν^*. **(e)** *Top*: A single pattern ν^* is trained multiple times while all other patterns are trained once. *Bottom*: the number of times that pattern ν^* must be repeated during training in order to obtain a classification error rate below a threshold p_θ during testing.

Input alignment and control transfer in the two-pathway model

In the previous section, we showed that, due to the reinforcement of the input-to-output mapping that comes about by slow Hebbian learning in the single-neuron model, input from the second pathway can contribute to driving the readout unit to its target state. We next investigated how the learning process in the two-pathway model could be interpreted at the population level. In this case, while each individual readout neuron performs the same binary classification task as before, the activity of the readout population is characterized by an N_z-dimensional vector $\mathbf{z}$. Interpreting the two input populations as cortical and thalamic inputs to striatum, we asked what are the principles that determine the relationship of this striatal population activity to its inputs, and how does this relationship evolve through learning. To address this, we defined $\mathbf{m}$ and $\mathbf{h}$ as the N_z-dimensional inputs to the readout population from the first and second pathway, respectively (Figure 3a), where $m_i = \sum_{j=1}^{N_x} W_{ij} x_j$ and $h_i = \sum_{j=1}^{N_y} V_{ij} y_j$, and the weight vectors $\mathbf{w}$ and $\mathbf{v}$ from the single-neuron model have been promoted to weight matrices $\mathbf{W}$ and $\mathbf{V}$. The activity in the readout population is then described by the N_z-dimensional vector $\mathbf{z} = \mathrm{sgn}(\mathbf{m} + \mathbf{h})$. In this case, applying the updates (1) and (3) for each synapse when training on pattern μ leads to updated input currents $\mathbf{m}^\mu + \Delta\mathbf{m}^\mu = \hat{\mathbf{z}}^\mu - \mathbf{h}^\mu$ and $\mathbf{h}^\mu + \Delta\mathbf{h}^\mu = (1 - \alpha/N_y)\mathbf{h}^\mu + \sqrt{2}\beta n_\mu \hat{\mathbf{z}}^\mu/\bar{n}$ in the two pathways. Thus, following the updates, both of the input currents obtain a component of magnitude $\sim O(1)$ along the target activity $\hat{\mathbf{z}}^\mu$, so that they become aligned with the target activity and with each other. Further, for the Hebbian update, this component continues to grow as the pattern is repeated n_μ times in succession (Figure 3b).

As a single input pattern is repeated more times in succession, the normalized alignment between the inputs, defined as $\mathbf{m} \cdot \mathbf{h}/(|\mathbf{m}||\mathbf{h}|)$ grows, meaning that the inputs from the two pathways are driving the downstream population in similar ways (Figure 3c, top), a process that we refer to as *input alignment*. In addition, as the number of repetitions grows, the input from the second pathway along $\hat{\mathbf{z}}$ constitutes an increasingly dominant proportion of the total input along $\hat{\mathbf{z}}$ (Figure 3c, bottom), a process that we refer to as *control transfer*. Both of these processes—alignment of the two inputs and the transfer of control from one pathway to the other—are *covert*, in the sense that they occur gradually during repetition and do not have any immediately obvious behavioral correlate, since the learning in the first pathway causes the activity in the downstream population to attain its target value already by the first repetition.

In order to further illustrate the implications of input alignment and control transfer, we sequentially trained a two-pathway network with a multi-neuron readout population to produce P classifications, generalizing the task from previous sections to the case $N_z > 1$. During training, one particular pattern was repeated many times in succession, while all others were trained only once (Figure 3d, top). We found that input alignment became large and positive during training and, particularly for the repeated pattern, remained large and positive during testing after training was complete (Figure 3d, second panel). Similarly, the ratio of the input $\mathbf{h}$ along $\hat{\mathbf{z}}$ from the second pathway to the total input along $\hat{\mathbf{z}}$ became large for the repeated pattern during training and remained large during testing, illustrating the occurrence of control transfer (Figure 3d, third panel). Finally, when testing the error rate for each pattern after training was complete, we performed "lesions" by shutting off each of the inputs during testing (Figure 3d, bottom). In the case $\mathbf{m} = 0$, we observed that the error rate increased for all patterns *except* for the pattern that was repeated multiple times during training, which retained an error rate of zero. This shows that, via the mechanisms of input alignment and control transfer for the highly practiced behavior, the second pathway was able to compensate for the loss of the first pathway, driving the downstream population to produce the correct activity even in the absence of the first input. In the case $\mathbf{h} = 0$, on the other hand, the repeated pattern was not recalled accurately, illustrating that the input from the second pathway is necessary to protect practiced patterns from being overwritten. Thus, interpreting the two input pathways as cortical and thalamic inputs to striatum, the model is able to account for the experimental observations in rodents

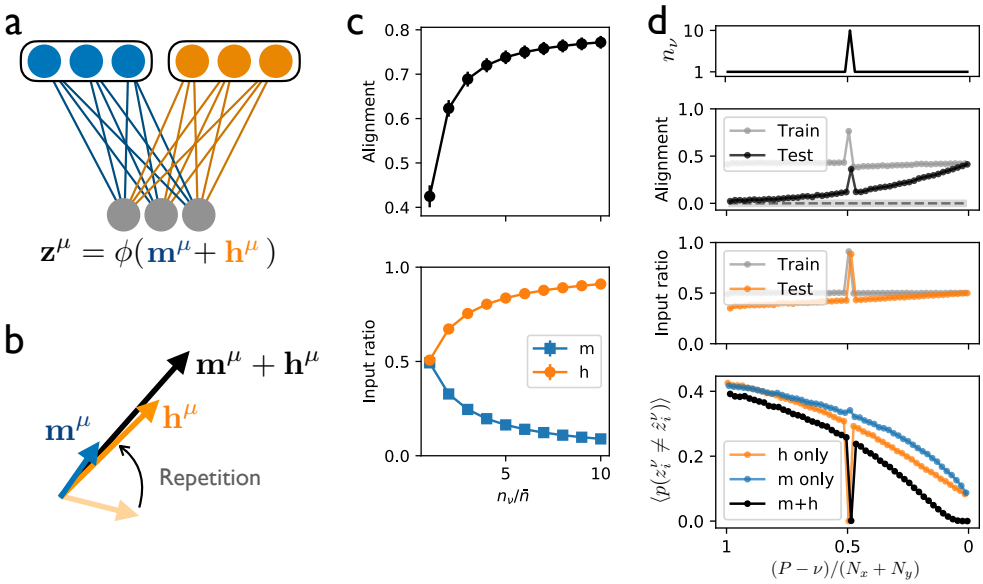


Figure 3: Inputs to the downstream population become aligned through repetition. **(a)** A downstream population receives input from two pathways, where the first input $\mathbf{m}$ is trained with fast supervised learning, while the second input $\mathbf{h}$ is modified with slow Hebbian learning. **(b)** Through repetition, the second input becomes increasingly aligned with the first, while its projection along the readout direction $\hat{\mathbf{z}}$ grows. **(c)** *Top*: The normalized overlap between $\mathbf{m}$ and $\mathbf{h}$ for one particular pattern that is repeated n_ν times. *Bottom*: The ratio of the projection of each input component along the readout direction ($\mathbf{m}\cdot\hat{\mathbf{z}}$ or $\mathbf{h}\cdot\hat{\mathbf{z}}$) to the total input along the readout direction ($(\mathbf{m}+\mathbf{h})\cdot\hat{\mathbf{z}}$). **(d)** *Top*: In a two-pathway network trained on P patterns sequentially, one pattern is repeated 10 times during training, while others are trained only once. *Middle panels*: The alignment for each pattern during sequential training and testing (upper panel, where shaded region is chance alignment for randomly oriented vectors), and the ratio of the input $\mathbf{h}$ along $\hat{\mathbf{z}}$ to the total input along $\hat{\mathbf{z}}$ (lower panel). *Bottom*: The error rate, averaged over N_z readout units, for the trained network with either both inputs intact (black) or with one input removed (blue and orange) (curves offset slightly for clarity). All points in (c)-(d) are simulations with $\alpha = \beta = 1$ and $N_x = N_y = N_z = 1000$.

that skilled behaviors and neural representations of kinematic variables are robust to the loss of cortical inputs [28, 30, 29], but that such behaviors are not robust to the loss of thalamic input [27].

In summary, when an input pathway with fast supervised learning and another with slow associative learning together drive a downstream population, the mechanisms of input alignment and control transfer enable the second input to gradually take over the role of the first in driving the downstream activity. For highly practiced patterns, this enables the second input to produce the correct downstream activity even if the first input is removed completely. As shown in Supplemental Figure 6, these mechanisms also occur in the case where reinforcement learning is used rather than supervised learning in the first pathway, but not in the case where both pathways feature reward-driven learning, illustrating that the combination of Hebbian learning together with error- or reward-based learning is necessary in order for input alignment and control transfer to take place.

Implementation in the sensorimotor brain circuit: Automatization of behavior through practice

In the previous section, we showed that, for a single-layer neural network in which two input pathways operate with different types of learning, fast error-driven learning in the first pathway causes the downstream neurons to produce the correct output, while slow associative learning in the second pathway reinforces the association between the input to the network and the correct output. After many repetitions, this scheme enables the second input pathway to assume control for driving the downstream layer. In this section, we develop the experimental interpretation of the model in greater detail, using it to address existing experimental results and to formulate predictions for future experiments.

As a neurobiological implementation of the model, we proposed identifying the readout population in the two-pathway model as sensorimotor striatum, the input population with fast supervised learning as motor cortex, and the input population with slow associative learning as thalamus. Taken together, these identifications suggest a picture in which information about state, including sensory feedback, cues, and context, is represented in cortex and thalamus, and in which plasticity at the synapses descending from these structures to striatum allows for state information to be mapped onto actions that will tend to minimize errors and maximize rewards (Figure 4a). The mechanisms and roles of learning in the two pathways are distinct, with fast changes to corticostriatal synapses serving to learn and produce the correct behavior, while slower changes to thalamostriatal synapses serve to associate a given state with a behavioral output, with this association becoming stronger and more persistent the more that it is practiced. As described in the previous section, the transfer of control from corticostriatal to thalamostriatal inputs would suggest that motor cortex should be necessary for the initial learning of a behavior, but not for the production of the behavior after a sufficient amount of practice. This is consistent with lesion and inactivation studies in rodents [28, 30], as well as with studies of learned motor behaviors in humans suggesting that motor cortex becomes less involved in driving behaviors with practice [31, 32].

With the neurobiological implementation proposed above, the two-pathway model also leads to a number of testable experimental predictions. First and most obviously, the theory predicts that cortico- vs. thalamostriatal synapses should be modified by different types of plasticity (fast and reward-driven vs. slow and associative, respectively). Because both synapses are glutamatergic, they have typically not been studied separately or even distinguished in past experiments [47]. Future experiments could test for differences between the types of modifications at these synapses, either *in vivo* or in slice preparations.

Second, because motor cortex plays less of a role in driving striatum as behaviors become highly practiced, lesioning, inactivating, or randomly perturbing activity in motor cortex during behavior should have a

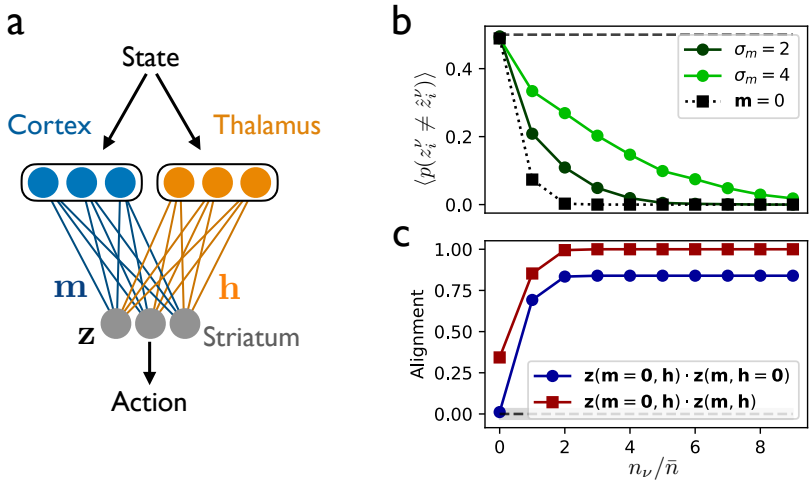


Figure 4: Two-pathway implementation in the sensorimotor brain circuit and predicted responses to perturbations. **(a)** In the proposed neurobiological implementation of the two-pathway model, motor cortex and thalamus provide inputs to the sensorimotor region of striatum. **(b)** The population-averaged error rate in the readout population in the presence of noise perturbations to $\mathbf{m}$ (noise amplitude σ_m, green lines) or upon removing the input $\mathbf{m}$ (dotted line) for a single pattern repeated n times during training. **(c)** The normalized alignment between the population activity $\mathbf{z}$ with $\mathbf{m} = \mathbf{0}$ and with $\mathbf{h} = \mathbf{0}$ (blue line), and the normalized alignment between $\mathbf{z}$ with $\mathbf{m} = \mathbf{0}$ and with $\mathbf{m} \neq \mathbf{0}$. In (b) and (c), points are from simulations in a network with $\alpha = \beta = 1$ and $N_x = N_y = N_z = 1000$.

decreasingly small effect on a learned behavior the more that it has been trained (Figure 4b). Furthermore, due to the fact that covert learning in the second pathway should occur slowly relative to learning in the first pathway, the learned behavior should become robust to motor-cortex inactivation only some time after the behavior has been mastered. This effect has in fact already been observed in very recent experiments using optogenetic inactivation throughout learning of a targeted forelimb task [30].

Third, online optogenetic inactivation could be used so that only motor cortex or only thalamus drives striatum in a given trial. The theory would predict that the striatal activities observed in these two cases would be correlated due to input alignment, with the degree of correlation increasing with the amount of training (Figure 4c, blue line). Similarly, striatal population activity should be similar before and after cortical lesion or inactivation in a trained animal, with the degree of similarity increasing with the amount of training (Figure 4c, red line).

Finally, if thalamostriatal plasticity is blocked during training of a behavior, then the model predicts that the behavior will still be learned due to the fact that motor cortex is responsible for driving initial learning. Since input alignment cannot occur in the absence of thalamostriatal plasticity, however, the learned behavior will no longer be robust to cortical lesion or inactivation. This would be a challenging experiment to perform due to the fact that both thalamo- and corticostriatal synapses are glutamatergic, making selective manipulation difficult. If it could be done, however, it would provide strong evidence for our proposed neurobiological implementation of the two-pathway model.

Simulated reaching task with the two-pathway network

With the experimental interpretation from the preceding section in mind, we asked whether the basic mechanism illustrated in Figure 2—namely, the enhanced retention of learned classifications that are practiced multiple times—could also be implemented in a neural network performing a more challenging sensorimotor

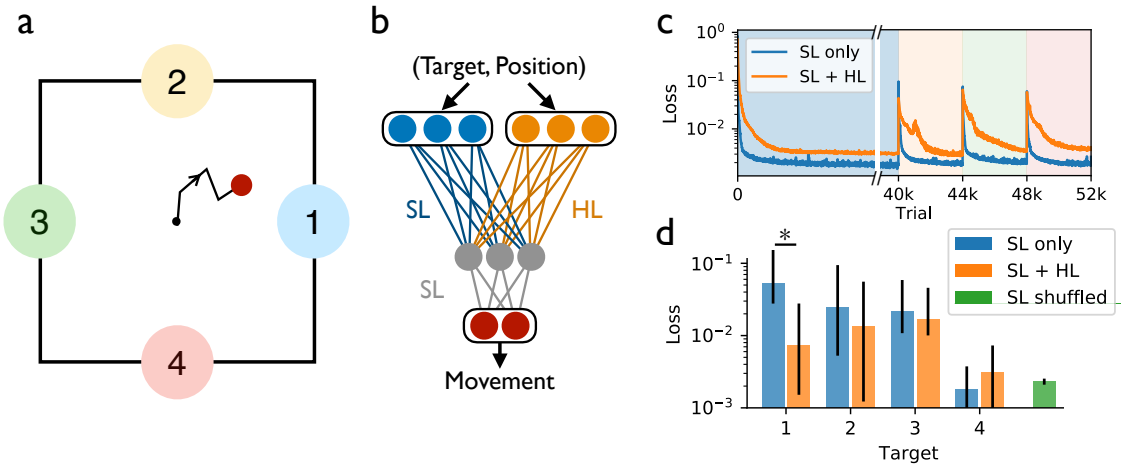


Figure 5: A two-pathway neural network trained to perform a reaching task exhibits enhanced retention of highly practiced behaviors. **(a)** In this task, the agent, given one of four cues, is trained to perform a center-out reach over 10 timesteps to a particular target. **(b)** The neural network trained to perform the task receives a target-specific cue and the current position as input. The network contains one pathway learned with supervised learning (SL) and another with Hebbian learning (HL). The output is a two-dimensional vector determining the movement velocity in the next timestep. **(c)** The loss (mean squared error) on each of the four targets in sequence during training, with a longer training period for the first target, and with Hebbian learning either active (SL+HL) or inactive (SL only). **(d)** The loss during testing on each of the four targets, after training is complete. The green bar shows a control network in which targets were randomly interleaved during training.

task. We thus extended the two-pathway architecture to a neural network with one hidden layer and trained the network to perform a center-out reaching task (Figure 5a,b). In this task, the velocity of a cursor moving in two dimensions is controlled by a neural network, which receives as input the current location of the cursor together with a context input indicating which target to move toward. In each trial, the network was trained with supervised learning to follow a smooth, minimal-jerk trajectory to the specified target, with each trial having a duration of 10 timesteps. As in the single-neuron two-pathway model, the inputs were divided into two pathways, with supervised learning in one pathway and Hebbian learning in the other.

The network was first trained in a block of trials to reach to target 1, then in another block to reach to target 2, and so on, with more trials in the first block than in the later blocks (Figure 5). When the network was tested after training was complete, we found that, when Hebbian learning was activated, the performance for target 1 was comparable to or even better than that for target 2, despite the fact that the training for target 2 had occurred more recently (Figure 5d). This did not occur in the network without Hebbian learning, where performance for target 1 was worse than that for other targets despite the larger amount of practice ($p = 0.011$). This result shows that, in this trained network model as in the single-neuron case, associative Hebbian learning enables practiced behaviors to be better retained, protecting them from being overwritten as new behaviors are learned.

Discussion

In this work, we have developed a bottom-up, mechanistic theory of fast and slow learning in parallel descending pathways, applying this theory to obtain insight into the possible roles of cortical and thalamic inputs to the basal ganglia in the acquisition and retention of learned behaviors through practice. While

two-pathway architectures have been proposed previously as a means of addressing multistage learning in mammalian [48, 49] and songbird [50] motor circuits, those models featured different learning mechanisms and were applied to different anatomical structures than the two-pathway model studied here. In particular, our theory proposes two simultaneously occurring learning processes: a fast process that minimizes errors in the cortical pathway, paired with a slow process that strengthens associations in the thalamic pathway. From a computational perspective, this work provides a quantitative theoretical framework in which the effects of learning, forgetting, and practice can be analyzed for the simplest possible case, namely a single neuron classifying random input patterns. From a cognitive perspective, the work provides an explanation for why repeated practice can continue to have effects even after a memory or behavior has been learned to be performed with zero error. From a neurobiological perspective, this work provides a framework for understanding the automatization of learned behavior and the formation of habits, proposing how these processes might be implemented in neural circuits.

In the Introduction, we described the multiple effects of practice: improved performance, habit formation, and decrease in cognitive effort [3]. The two-pathway model—and its proposed implementation in the brain's sensorimotor circuitry in particular—provides a unified account of these effects. First, fast learning in the corticostriatal pathway provides a mechanism for improved performance, with the greatest gains coming in the early stages of practice. Second, the two-pathway model provides a mechanistic description of habit formation, in which behaviors are produced consistently in response to particular stimuli, even when the behavior is no longer rewarded. This is because information about rewards and errors is encoded in the updates to corticostriatal synapses, which play less and less of a role in driving downstream activity the more that a behavior is practiced. Going further, the model may even provide a description of addiction and compulsion, in which behaviors persist even if they lead to negative outcomes. Finally, the two-pathway model provides a mechanistic description of behavioral automatization, with the amount of conscious effort devoted to a behavior decreasing with increasing amounts of practice, as control of the behavior is transferred from a cortical to a subcortical circuit.

While the model that we have developed was constructed to address procedural learning of motor behaviors, the results shown in Figure 2c-e suggest that it may also lead to insights for declarative memory. Not all of the results obtained from the two-pathway model were in agreement with published results from experiments on human memory, however. The fact that memory recall performance decays roughly as a power law has been relatively well established experimentally [45]. However, the forgetting curves in the two-pathway model are better described by exponential decay than by power laws (Supplemental Figure 4). In addition, experimental study of spaced-repetition effects in human memory has shown that there exists an optimal interval between presentations during training [51], such that the interval between the two training presentations should match that between the second presentation and the time of testing. In the two-pathway model, however, we found that shorter training intervals always lead to better testing performance, regardless of the testing interval (Supplemental Figure 5). Of course, it is not surprising that a simple single-neuron model fails to capture the full range of experimental results on human memory. Characterizing the additional mechanisms that could lead to such effects as power law memory decay and optimal intervals for spaced repetition will be a promising direction for future study.

A promising direction for future work would be to extend to two-pathway model in ways that might enable it to address effects such as power-law memory decay and spaced repetition. For example, one successful approach for obtaining power law memory decay has been to allow individual synapses to have multiple states, with a different timescale associated with each state [52, 39], and incorporating associative learning and the effects of practice into such multi-state models could be a promising direction for future work on the mechanistic modeling of learning and memory. More broadly, future work in this area will need to identify

the network-level mechanisms that emerge in circuits with large numbers of neurons and multiple layers and might underlie phenomena such as power-law forgetting curves and the effects of spaced repetition. In addition, memory performance can also be affected by higher-level cognitive strategies such as chunking or creating narratives to link information meaningfully [45]. Clearly, such complex strategies cannot be addressed by a single-neuron model, though it might be hoped that such a model could be a building block for addressing such phenomena in the future.

Future work should also extend the model from supervised learning to reinforcement learning, which is less theoretically tractable but more plausible as a model of reward-based learning at corticostriatal synapses [14]. Reinforcement learning is also more plausible from a behavioral perspective, since it allows for performance to improve gradually by trial and error, leading to a richer description of learned motor behaviors than the single-step perceptron learning rule can provide. While a first step in this direction was taken in Supplemental Figure 6, which shows that the mechanisms of input alignment and control transfer are realized in a circuit trained with reinforcement learning, extending the theoretical framework underlying the two-pathway model to the case of reinforcement learning is an important direction for future work. Finally, in this work we have focused on sensorimotor striatum and its inputs. The thalamocortico-basal ganglia circuits that are involved in higher-level cognitive functions and that exist in parallel with the sensorimotor circuit could provide alternative interpretations of the model that we have presented [12, 53]. More broadly, this two-pathway motif might be found in other brain areas and could in part explain the large degree of apparent redundancy found throughout the neural circuitry of the brain.

Acknowledgments

The authors are grateful to B.P. Ölveczky, members of the Ölveczky lab, and L.F. Abbott for helpful discussions and feedback. Support for this work was provided by the National Science Foundation NeuroNex program (DBI-1707398), the National Institutes of Health (R01 NS105349, DP5 OD019897, and U19 NS104649), the Leon Levy Foundation, and the Gatsby Charitable Foundation.

References

[1] Gordon D Logan. Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 39(2):367, 1985.

[2] Agnes Moors and Jan De Houwer. Automaticity: a theoretical and conceptual analysis. *Psychological bulletin*, 132(2):297, 2006.

[3] Adrian M Haith and John W Krakauer. The multiple effects of practice: skill, habit and reduced cognitive load. *Current Opinion in Behavioral Sciences*, 20:196–201, 2018.

[4] John W Krakauer, Alkis M Hadjiosif, Jing Xu, Aaron L Wong, and Adrian M Haith. Motor learning. *Comprehensive Physiology*, 9:613–663, 2019.

[5] Anthony Dickinson. Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 308(1135):67–78, 1985.

[6] Ray J Dolan and Peter Dayan. Goals and habits in the brain. *Neuron*, 80(2):312–325, 2013.

[7] Gabriele Wulf. *Attention and motor skill learning*. Human Kinetics, 2007.

[8] Eran Dayan and Leonardo G Cohen. Neuroplasticity subserving motor skill learning. *Neuron*, 72(3):443–454, 2011.

[9] Hajime Mushiake and Peter L Strick. Pallidal neuron activity during sequential arm movements. *Journal of Neurophysiology*, 74(6):2754–2758, 1995.

[10] Shigehiro Miyachi, Okihide Hikosaka, Karadi Miyashita, Zoltán Kárádi, and Miya Kato Rand. Differential roles of monkey striatum in learning of sequential hand movement. *Experimental Brain Research*, 115(1):1–5, 1997.

[11] Okihide Hikosaka, Yoriko Takikawa, and Reiko Kawagoe. Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiological reviews*, 80(3):953–978, 2000.

[12] Henry H Yin and Barbara J Knowlton. The role of the basal ganglia in habit formation. *Nature Reviews Neuroscience*, 7(6):464, 2006.

[13] Ann M Graybiel. Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.*, 31:359–387, 2008.

[14] P Redgrave, N Vautrelle, and JNJ Reynolds. Functional properties of the basal ganglia's re-entrant loop architecture: selection and reinforcement. *Neuroscience*, 198:138–151, 2011.

[15] Sten Grillner and Brita Robertson. The basal ganglia downstream control of brainstem motor centres?an evolutionarily conserved strategy. *Current opinion in neurobiology*, 33:47–52, 2015.

[16] José L Lanciego, Natasha Luquin, and José A Obeso. Functional neuroanatomy of the basal ganglia. *Cold Spring Harbor perspectives in medicine*, 2(12):a009621, 2012.

[17] Barbara J Hunnicutt, Bart C Jongbloets, William T Birdsong, Katrina J Gertz, Haining Zhong, and Tianyi Mao. A comprehensive excitatory input map of the striatum reveals novel functional organization. *Elife*, 5:e19103, 2016.

[18] Houri Hintiryan, Nicholas N Foster, Ian Bowman, Maxwell Bay, Monica Y Song, Lin Gou, Seita Yamashita, Michael S Bienkowski, Brian Zingg, Muye Zhu, et al. The mouse cortico-striatal projectome. *Nature Neuroscience*, 19(8):1100, 2016.

[19] Okihide Hikosaka, Hiroyuki Nakahara, Miya K Rand, Katsuyuki Sakai, Xiaofeng Lu, Kae Nakamura, Shigehiro Miyachi, and Kenji Doya. Parallel neural networks for learning sequential procedures. *Trends in neurosciences*, 22(10):464–471, 1999.

[20] Patrick E Rothwell, Scott J Hayton, Gordon L Sun, Marc V Fuccillo, Byung Kook Lim, and Robert C Malenka. Input-and output-specific regulation of serial order performance by corticostriatal circuits. *Neuron*, 88(2):345–356, 2015.

[21] Yoland Smith, Dinesh V Raju, Jean-Francois Pare, and Mamadou Sidibe. The thalamostriatal system: a highly specific network of the basal ganglia circuitry. *Trends in neurosciences*, 27(9):520–527, 2004.

[22] Gil Mandelbaum, Julian Taranda, Trevor M Haynes, Daniel R Hochbaum, Kee Wui Huang, Minsuk Hyun, Kannan Umadevi Venkataraju, Christoph Straub, Wengang Wang, Keiramarie Robertson, et al. Distinct cortical-thalamic-striatal circuits through the parafascicular nucleus. *Neuron*, 102(3):636–652, 2019.

[23] Naoyuki Matsumoto, Takafumi Minamimoto, Ann M Graybiel, and Minoru Kimura. Neurons in the thalamic CM-Pf complex supply striatal neurons with information about behaviorally significant sensory events. *Journal of neurophysiology*, 85(2):960–976, 2001.

[24] Edgar Díaz-Hernández, Rubén Contreras-López, Asai Sánchez-Fuentes, Luis Rodríguez-Sibrían, Josué O Ramírez-Jarquín, and Fatuel Tecuapetla. The thalamostriatal projections contribute to the initiation and execution of a sequence of movements. *Neuron*, 100(3):739–752, 2018.

[25] Shigeki Kato, Ryoji Fukabori, Kayo Nishizawa, Kana Okada, Nozomu Yoshioka, Masateru Sugawara, Yuko Maejima, Kenju Shimomura, Masahiro Okamoto, Satoshi Eifuku, et al. Action selection and flexible switching controlled by the intralaminar thalamic neurons. *Cell reports*, 22(9):2370–2382, 2018.

[26] Laura Anne Bradfield, Genevra Hart, and Bernard W Balleine. The role of the anterior, mediodorsal, and parafascicular thalamus in instrumental conditioning. *Frontiers in systems neuroscience*, 7:51, 2013.

[27] Steffen B. E. Wolff, Raymond Ko, and Bence P. Ölveczky. Distinct roles for motor cortical and thalamic inputs to striatum during motor learning and execution. *bioRxiv*, 2019.

[28] Risa Kawai, Timothy Markman, Rajesh Poddar, Raymond Ko, Antoniu L Fantana, Ashesh K Dhawale, Adam R Kampff, and Bence P Ölveczky. Motor cortex is required for learning but not for executing a motor skill. *Neuron*, 86(3):800–812, 2015.

[29] Ashesh K. Dhawale, Steffen B. E. Wolff, Raymond Ko, and Bence P. Ölveczky. The basal ganglia can control learned motor sequences independently of motor cortex. *bioRxiv*, 2019.

[30] Eun Jung Hwang, Jeffrey E Dahlen, Yvonne Yuling Hu, Karina Aguilar, Bin Yu, Madan Mukundan, Akinori Mitani, and Takaki Komiyama. Disengagement of motor cortex from movement control during long-term learning. *Science Advances*, 5(10):eaay0001, 2019.

[31] Timo Krings, Rudolf Töpper, Henrik Foltys, Stephan Erberich, Roland Sparing, Klaus Willmes, and Armin Thron. Cortical activation patterns during complex motor tasks in piano players and control subjects. a functional magnetic resonance imaging study. *Neuroscience Letters*, 278(3):189–193, 2000.

[32] Liangsuo Ma, Binquan Wang, Shalini Narayana, Eliot Hazeltine, Xiying Chen, Donald A Robin, Peter T Fox, and Jinhu Xiong. Changes in regional activity are accompanied with changes in inter-regional connectivity during 4 weeks motor learning. *Brain Research*, 1318:64–76, 2010.

[33] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133, 1943.

[34] Koby Crammer, Ofer Dekel, Joseph Keshet, Shai Shalev-Shwartz, and Yoram Singer. Online passive-aggressive algorithms. *Journal of Machine Learning Research*, 7(Mar):551–585, 2006.

[35] Thomas M Cover. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE transactions on electronic computers*, (3):326–334, 1965.

[36] Elizabeth Gardner. The space of interactions in neural network models. *Journal of Physics A*, 21(1):257, 1988.

[37] Michael McCloskey and Neal J Cohen. Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of Learning and Motivation*, volume 24, pages 109–165. Elsevier, 1989.

[38] Stefano Fusi and LF Abbott. Limits on the memory storage capacity of bounded synapses. *Nature neuroscience*, 10(4):485, 2007.

[39] Marcus K Benna and Stefano Fusi. Computational principles of synaptic memory consolidation. *Nature neuroscience*, 19(12):1697, 2016.

[40] Donald Olding Hebb. *The Organization of Behavior: A Neuropsychological Theory*. 1949.

[41] William Charles Frederick Krueger. The effect of overlearning on retention. *Journal of Experimental Psychology*, 12(1):71, 1929.

[42] Benton J Underwood and Geoffrey Keppel. Retention as a function of degree of learning and letter-sequence interference. *Psychological Monographs: General and Applied*, 77(4):1, 1963.

[43] Leo Postman and Donald A Riley. Degree of learning and interserial interference in retention: A review of the literature and an experimental analysis. *University of California Publications in Psychology*, 1959.

[44] Norman J Slamecka and Brian McElree. Normal forgetting of verbal lists as a function of their degree of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(3):384, 1983.

[45] John Robert Anderson. *Learning and Memory: An Integrated Approach*. John Wiley & Sons Inc, 2000.

[46] John R Anderson. Interference: The relationship between response latency and response accuracy. *Journal of Experimental Psychology: Human Learning and Memory*, 7(5):326, 1981.

[47] Charles R Gerfen and D James Surmeier. Modulation of striatal projection systems by dopamine. *Annual review of neuroscience*, 34:441–466, 2011.

[48] Hiroshi Makino, Eun Jung Hwang, Nathan G Hedrick, and Takaki Komiyama. Circuit mechanisms of sensorimotor learning. *Neuron*, 92(4):705–721, 2016.

[49] Ryan Pyle and Robert Rosenbaum. A reservoir computing model of reward-modulated motor learning and automaticity. *Neural computation*, 31(7):1430–1461, 2019.

[50] Tiberiu Teşileanu, Bence Ölveczky, and Vijay Balasubramanian. Rules and mechanisms for efficient two-stage learning in neural circuits. *Elife*, 6:e20944, 2017.

[51] Arthur M Glenberg. Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior*, 15(1):1–16, 1976.

[52] Stefano Fusi, Patrick J Drew, and Larry F Abbott. Cascade models of synaptically stored memories. *Neuron*, 45(4):599–611, 2005.

[53] Marjan Jahanshahi, Ignacio Obeso, John C Rothwell, and José A Obeso. A fronto–striato–subthalamic–pallidal network for goal-directed and habitual inhibition. *Nature Reviews Neuroscience*, 16(12):719, 2015.

[54] H. Risken. *Fokker-Planck Equation*. Springer, 1996.

[55] Ronald J Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3-4):229–256, 1992.

[56] Richard S Sutton and Andrew G Barto. *Reinforcement Learning: An Introduction*. MIT press, 2018.

Materials and Methods

Parameters and equations for the simulations in Figures 1, 2, 3, and 4, as well as Supplemental Figures 2 and 3, are provided in the Main Text.

For the simulations in Figure 5, $n = 21$ networks were trained for each of the conditions shown (SL only, SL+HL, and SL shuffled), where each network had $N_x = N_y = 50$, $N_z = 10$, and a 2-dimensional linear readout. The readout controlled the velocity of a cursor, which began at position $\mathbf{r} = \mathbf{0}$ and moved to one of the four target positions: (1,0), (0,1), (-1,0), (0,-1). The target velocity in each trial was given by

$$\hat{\mathbf{u}}(t) = \mathbf{r}^i \left[30(t/T)^2 - 60(t/T)^3 + 30(t/T)^4\right],$$

where t is the timestep, $T = 10$ is the number of timesteps in each trial, and $\mathbf{r}^i$ denotes one of the four target positions. This bell-shaped velocity profile describes a minimum-jerk trajectory from the initial to the target point. Both input layer populations receive fixed random projections of the network readout from the previous timestep, as well as a tonic random input associated with one of the four targets.

During training, the mean squared difference between the network output and $\hat{\mathbf{u}}(t)$ was minimized by gradient descent on the weights $\mathbf{W}$ (labelled SL in Figure 5b), with learning rate $\eta_{\mathrm{SL}} = 0.001$ (found by grid search). The weights $\mathbf{V}$ (labelled HL in Figure 5b) were either not trained (in the cases SL only and SL shuffled) or trained with the Hebbian update rule $\Delta V_{ij} = \eta_{\mathrm{HL}}[-V_{ij} + z_i(t)y_j(t)]$, where $\eta_{\mathrm{HL}} = 10^{-6}$ was found by grid search. The networks were sequentially trained on each target, as shown in Figure 5c. The bars in Figure 5d show the median test loss over the trained networks for each target, with error bars denoting the 25th and 75th percentiles.

Appendix 1: The perceptron forgetting curve

As described in the main text, we wish to train a perceptron having $N_x = N$ inputs and subject to the update rule (1) to map P random input patterns onto randomly chosen binary outputs. The question that we seek to answer is, after P patterns have been trained using the update rule in (1), what will be the probability of misclassification if we then test the output produced by a particular pattern ν without any further learning? Clearly, the most recently learned patterns are likely to produce correct outputs, while those learned long in the past are more likely to produce errors due to accumulated changes in the weights $\mathbf{w}$ during subsequent learning. In general, the probability of an error when testing on pattern ν is, using the Heaviside step function $\Theta(\cdot)$, given by

$$\begin{aligned} p(z^\nu \neq 1) = \int d\mathbf{x}^\nu p(\mathbf{x}^\nu) \int d\mathbf{w}^\nu p(\mathbf{w}^\nu) \int d\mathbf{w}^P \Big[& p(\mathbf{w}^P|\mathbf{w}^\nu + [1-\mathbf{w}^\nu\cdot\mathbf{x}^\nu]\mathbf{x}^\nu/N)\Theta(-\mathbf{w}^P\cdot\mathbf{x}^\nu)\Theta(1-\mathbf{w}^\nu\cdot\mathbf{x}^\nu) \\ & + p(\mathbf{w}^P|\mathbf{w}^\nu)\Theta(-\mathbf{w}^P\cdot\mathbf{x}^\nu)\Theta(\mathbf{w}^\nu\cdot\mathbf{x}^\nu - 1)\Big]. \end{aligned} \tag{5}$$

In this equation, we have, without loss of generality, redefined $\mathbf{x}^\mu \to \hat{z}^\mu \mathbf{x}^\mu$, so that the target output becomes $z^\mu = 1$ for every pattern. We have also set the classification margin $\kappa = 1$, which amounts to a choice for scaling the overall magnitude $|\mathbf{w}|$. In this equation, the weight vector just before training pattern ν is assumed to come from a distribution $p(\mathbf{w}^\nu)$, which we shall derive below. The first line in (5) counts the cases in which the classification using this weight vector is initially incorrect or correct with margin less than $\kappa = 1$ (so $1 - \mathbf{w}^\nu \cdot \mathbf{x} > 0$) and in which, after making the initial update $\mathbf{w}^\nu \to \mathbf{w}^\nu + \Delta\mathbf{w}^\nu$, the weight vector evolves through $P - \nu$ successive updates into the final weight vector $\mathbf{w}^P$, which leads to incorrect classification when pattern ν is again tested ($-\mathbf{w}^P \cdot \mathbf{x}^\nu > 0$). Similarly, the second line of (5) counts the cases in which the classification is initially *correct* with a sufficiently large margin and in which, after making successive weight updates, the final weight vector $\mathbf{w}^P$ again leads to incorrect classification.

At this stage, the probability distributions $p(\mathbf{w}^\nu)$ and $p(\mathbf{w}^P|\mathbf{w}^\nu)$ in (5) are unknown. For the latter distribution, however, we can track its evolution step by step using the update rule (1). Because $\mathbf{x}$ is a random variable, we first seek to find the distribution $p(\Delta\mathbf{w}|\mathbf{w})$ by averaging over $\mathbf{x}$. In fact, it will be sufficient just to calculate the first two moments of this distribution. Let $\mathbf{x} = \mathbf{x}^\parallel + \mathbf{x}^\perp$, where $\mathbf{x}^\parallel$ is the component along $\mathbf{w}$, and the index ν has been dropped for simplicity. In this case the weight update (1), in the case where an update occurs, can be written as

$$\Delta\mathbf{w} = \frac{1}{N}(1 - |\mathbf{w}|x^\parallel)(x^\parallel \mathbf{w}/|\mathbf{w}| + \mathbf{x}^\perp). \tag{6}$$

Using (6), in the case where a weight update occurs, the first moment is given by

$$\begin{aligned} \mu_i(\mathbf{w}) &\equiv \langle \Delta w_i \rangle_{\mathbf{x}} \\ &= \frac{1}{N}\int \frac{dx^\parallel}{\sqrt{2\pi}} e^{-(x^\parallel)^2/2} \int \frac{d\mathbf{x}^\perp}{(2\pi)^{(N-1)/2}} e^{-|\mathbf{x}^\perp|^2/2}(1 - |\mathbf{w}|x^\parallel)(x^\parallel w_i/|\mathbf{w}| + x_i^\perp) \\ &= -\frac{1}{N} w_i \int \frac{dx^\parallel}{\sqrt{2\pi}} e^{-(x^\parallel)^2/2}(x^\parallel)^2 \\ &= -\frac{1}{N} w_i. \end{aligned} \tag{7}$$

Similarly, the second moment of the distribution is

$$
\begin{aligned}
\Sigma_{ij}(\mathbf{w}) &\equiv \langle \Delta w_i \Delta w_j \rangle_{\mathbf{x}} - \mu_i(\mathbf{w})\mu_j(\mathbf{w}) \\
&= \frac{1}{N^2} \int \frac{dx^{\parallel}}{\sqrt{2\pi}} e^{-(x^{\parallel})^2/2} \int \frac{d\mathbf{x}^{\perp}}{(2\pi)^{(N-1)/2}} e^{-|\mathbf{x}^{\perp}|^2/2} (1 - |\mathbf{w}| x^{\parallel})^2 \left(\frac{x^{\parallel} w_i}{|\mathbf{w}|} + x_i^{\perp} \right) \left(\frac{x^{\parallel} w_j}{|\mathbf{w}|} + x_j^{\perp} \right) - \mu_i(\mathbf{w})\mu_j(\mathbf{w}) \\
&= \frac{1}{N^2} \int \frac{dx^{\parallel}}{\sqrt{2\pi}} e^{-(x^{\parallel})^2/2} \left[1 + |\mathbf{w}|^2 (x^{\parallel})^2 \right] \left[(x^{\parallel})^2 \frac{w_i w_j}{|\mathbf{w}|^2} + \delta_{ij} \right] - \mu_i(\mathbf{w})\mu_j(\mathbf{w}) \\
&= \frac{2}{N^2} \left[g(|\mathbf{w}|)\delta_{ij} + h(|\mathbf{w}|) w_i w_j \right]
\end{aligned}
\tag{8}
$$

where

$$
\begin{aligned}
g(|\mathbf{w}|) &\equiv \frac{1 + |\mathbf{w}|^2}{2} \\
h(|\mathbf{w}|) &\equiv \frac{2|\mathbf{w}|^2 + 1}{2|\mathbf{w}|^2}.
\end{aligned}
\tag{9}
$$

Together, (7)-(9) describe a drift-diffusion process. If we neglect higher-order moments, then the single-step probability distribution is given by

$$
p(\Delta \mathbf{w} | \mathbf{w}) = q \mathcal{N}(\boldsymbol{\mu}(\mathbf{w}), \boldsymbol{\Sigma}(\mathbf{w})) + (1 - q)\delta(\Delta \mathbf{w}), \tag{10}
$$

where q (to be calculated below) is the probability that a weight update occurs in a given step, and $\mathcal{N}(\cdot, \cdot)$ is the multinormal distribution. The time evolution of the probability distribution of the weights is then given by the Fokker-Planck equation [54]:

$$
\frac{dp(\mathbf{w}^{\tau})}{d\tau} = -\sum_i \frac{\partial}{\partial w_i^{\tau}} \left[\mu_i(\mathbf{w}^{\tau}) p(\mathbf{w}^{\tau}) \right] + \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial w_i^{\tau} \partial w_j^{\tau}} \left[\Sigma_{ij}(\mathbf{w}^{\tau}) p(\mathbf{w}^{\tau}) \right], \tag{11}
$$

where the initial condition from (5) is either $p(\mathbf{w}) = \delta(\mathbf{w} - \mathbf{w}^{\nu})$ or $p(\mathbf{w}) = \delta(\mathbf{w} - \mathbf{w}^{\nu} - \Delta \mathbf{w}^{\nu})$, and $\tau \equiv q(P - \nu)$ is the effective time variable.

Because the coefficients in (11) depend on $|\mathbf{w}^{\tau}|$, the full solution is not known in general. However, it is straightforward to calculate the time evolution of the moments of $p(\mathbf{w}^{\tau})$ by multiplying both sides by powers of $\mathbf{w}^{\tau}$ and using integration by parts [54]. Denoting the first two moments (not to be confused with the moments of $\Delta \mathbf{w}$ defined above) as

$$
\begin{aligned}
\tilde{\mu}_i(\tau) &\equiv \langle w_i \rangle \\
\tilde{\Sigma}_{ij}(\tau) &\equiv \langle w_i w_j \rangle - \langle w_i \rangle \langle w_j \rangle,
\end{aligned}
\tag{12}
$$

the time evolution is given by

$$
\begin{aligned}
\dot{\tilde{\mu}}_i &= -\frac{\tilde{\mu}_i}{N} \\
\dot{\tilde{\Sigma}}_{ij} &= -\frac{2}{N} \left(1 - \frac{h}{N} \right) \tilde{\Sigma}_{ij} + \frac{2h}{N^2} \tilde{\mu}_i \tilde{\mu}_j + \frac{2g}{N^2} \delta_{ij}.
\end{aligned}
\tag{13}
$$

In the $N \to \infty$ limit, we can assume (to be checked below) that $|\mathbf{w}^{\tau}|$ is constant. In this case, (13) has the solutions

$$
\begin{aligned}
\tilde{\mu}_i(\tau) &= -\frac{1}{N} w_i^{\nu} e^{-\tau/N} \\
\tilde{\Sigma}_{ij}(\tau) &\overset{(N \to \infty)}{=} \frac{2h\tau}{N^2} e^{-2\tau/N} w_i^{\nu} w_j^{\nu} + \delta_{ij} \frac{g}{N} \left(1 - e^{-2\tau/N} \right),
\end{aligned}
\tag{14}
$$

where terms $\sim O(1/N^2)$ in the second equation have been dropped in the large-N limit (while the first term is kept because τ may be $\sim O(N)$).

With (14), the solution to the Fokker-Planck equation (11) when just the first two moments are kept is

$$\begin{aligned} p(\mathbf{w}^P|\mathbf{w}^\nu) =& \frac{1}{[2\pi g(1-\gamma^2)/N]^{N/2}} \\ &\times \exp\left(-\frac{N}{2g(1-\gamma^2)}\left[\mathbf{w}^P - \gamma\mathbf{w}^\nu\right]\cdot\left[1 - \frac{2h\gamma^2 q(P-\nu)}{Ng(1-\gamma^2)}\mathbf{w}^\nu\otimes\mathbf{w}^\nu\right]\cdot\left[\mathbf{w}^P - \gamma\mathbf{w}^\nu\right]\right), \end{aligned} \tag{15}$$

where "$\otimes$" denotes the outer product (i.e. $[\mathbf{w}\otimes\mathbf{w}]_{ij} = w_i w_j$), and we have defined

$$\gamma \equiv e^{-q(P-\nu)/N}. \tag{16}$$

In the $N\to\infty$ limit, the anisotropic term can be ignored, giving

$$p(\mathbf{w}^P|\mathbf{w}^\nu) \overset{(N\to\infty)}{=} \frac{1}{[2\pi g(1-\gamma^2)/N]^{N/2}} \exp\left(-\frac{N}{2g(1-\gamma^2)}\left[\mathbf{w}^P - \gamma\mathbf{w}^\nu\right]^2\right), \tag{17}$$

which is the probability density evolution corresponding to an Ornstein-Uhlenbeck stochastic process [54]. After a long time, $\gamma\to 0$ and $p(\mathbf{w}^P|\mathbf{w}^\nu)$ from (15) approaches the steady-state distribution

$$p(\mathbf{w}) \equiv \frac{1}{[2\pi g/N]^{N/2}} \exp\left(-\frac{N}{2g}|\mathbf{w}|^2\right). \tag{18}$$

Thus, the deterministic update rule (1) leads to a bounded steady-state weight distribution $p(\mathbf{w})$ after a large number of classifications have been learned. This differs somewhat from previous models of sequential learning with random synaptic weight updates, such a bounded distribution as $P/N\to\infty$ was achieved either by requiring that the synaptic weights should be bounded [38] or that that they should decay slightly at each step [39].

Given the steady-state distribution (18), we can assume that $\hat{w}\equiv|\mathbf{w}|$ is constant in the large-N limit and let $\hat{g}\equiv g(\hat{w})$. Then, with $g=\hat{g}$, (18) can be used to calculate the variance of $\mathbf{w}$, leading to the self-consistent equation $\hat{w}^2=\hat{g}$, which, using the definition of $g(|\mathbf{w}|)$ from (9), has the solution $\hat{w}^2=\hat{g}=1$. This is close to but differs somewhat from the steady-state norm found in numerical simulations, from which $\hat{w}\simeq 1.19$. The reason for this is presumably because of the decision to approximate $p(\Delta\mathbf{w}|\mathbf{w})$ using only the first two moments of the distribution in (11). In general, the higher-order moments do not vanish, and these will contribute higher-order derivative terms in the Fokker-Planck equation (12). In turn, such terms will lead to nonvanishing higher-order moments in the distribution $p(\mathbf{w})$, beyond the two that were calculated in (12)-(14). Presumably, it is these higher-order terms which cause the discrepancy between the simulated result and the self-consistent calculation. In the theoretical curves shown in the Results section, we use the value of $\hat{w}$ obtained from simulations. Using (18), we can also calculate the probability of making a weight update in a given step, which is given by

$$\begin{aligned} q &= \int\frac{d\mathbf{w}}{(2\pi\hat{w}^2/N)^{N/2}}\int\frac{d\mathbf{x}}{(2\pi)^{N/2}}e^{-|\mathbf{x}|^2/2}e^{-N|\mathbf{w}|^2/2\hat{w}^2}\Theta(1-\mathbf{w}\cdot\mathbf{x}) \\ &= \frac{1}{2}\mathrm{erfc}\left(-\frac{1}{\sqrt{2}\hat{w}}\right) \end{aligned} \tag{19}$$

and evaluates to $q\simeq 0.798$.

With the preceding points in mind, and making use of the probability distributions (17) and (18), we can proceed to evaluate the integrals in (5) to obtain the probability of incorrect classification when testing

pattern ν. In order to factorize the arguments of the Heaviside step functions, we make use of the following identity:

$$\begin{aligned}\Theta(\rho) &= \int_0^\infty du\ \delta(u-\rho) \\ &= \int_0^\infty du \int_{-\infty}^\infty \frac{dv}{2\pi} e^{iv(u-\rho)}.\end{aligned} \tag{20}$$

with this trick, (5) becomes

$$p(z^\nu \neq 1) = I_1 + I_2, \tag{21}$$

where

$$\begin{aligned}I_1 = &\int_0^\infty du \int_0^\infty du' \int_{-\infty}^\infty \frac{dv}{2\pi} \int_{-\infty}^\infty \frac{dv'}{2\pi} e^{i(uv+u'v')} \int \frac{d\mathbf{x}}{(2\pi)^{N/2}} e^{-\mathbf{x}^2/2} \\ &\times \int \frac{d\mathbf{w}^\nu}{(2\pi\hat{g}/N)^{N/2}} e^{-N|\mathbf{w}^\nu|^2/2\hat{g}} \int \frac{d\mathbf{w}^P}{[2\pi(1-\gamma^2)\hat{g}/N]^{N/2}} e^{i[v\mathbf{w}^P\cdot\mathbf{x}+v'(\mathbf{w}^\nu\cdot\mathbf{x}-1)]} \\ &\times \exp\left(\frac{-N\left[\mathbf{w}^P - \gamma(\mathbf{w}^\nu + [1-\mathbf{w}^\nu\cdot\mathbf{x}]\mathbf{x}/N)\right]^2}{2\hat{g}(1-\gamma^2)}\right)\end{aligned} \tag{22}$$

and

$$\begin{aligned}I_2 = &\int_0^\infty du \int_0^\infty du' \int_{-\infty}^\infty \frac{dv}{2\pi} \int_{-\infty}^\infty \frac{dv'}{2\pi} e^{i(uv+u'v')} \int \frac{d\mathbf{x}}{(2\pi)^{N/2}} e^{-\mathbf{x}^2/2} \\ &\times \int \frac{d\mathbf{w}^\nu}{(2\pi\hat{g}/N)^{N/2}} e^{-N|\mathbf{w}^\nu|^2/2\hat{g}} \int \frac{d\mathbf{w}^P}{[2\pi(1-\gamma^2)\hat{g}/N]^{N/2}} e^{i[v\mathbf{w}^P\cdot\mathbf{x}+v'(1-\mathbf{w}^\nu\cdot\mathbf{x})]} \\ &\times \exp\left(\frac{-N\left[\mathbf{w}^P - \gamma\mathbf{w}^\nu\right]^2}{2\hat{g}(1-\gamma^2)}\right)\end{aligned} \tag{23}$$

Beginning with I_1, the integral over $\mathbf{w}^P$ can be performed, which, after simplification and using $\mathbf{x}^2 = N$ in the $N \to \infty$ limit, leads to

$$\begin{aligned}I_1 = &\int_0^\infty du \int_0^\infty du' \int_{-\infty}^\infty \frac{dv}{2\pi} \int_{-\infty}^\infty \frac{dv'}{2\pi} e^{i[(u+\gamma)v+(u'-1)v']} \int \frac{d\mathbf{x}}{(2\pi)^{N/2}} e^{-\mathbf{x}^2/2} \\ &\times \int \frac{d\mathbf{w}^\nu}{(2\pi\hat{g}/N)^{N/2}} \exp\left(-\frac{N}{2\hat{g}}|\mathbf{w}^\nu|^2 + iv'\mathbf{w}^\nu\cdot\mathbf{x} - \frac{\hat{g}(1-\gamma^2)}{2}v^2\right).\end{aligned} \tag{24}$$

Performing the integrals over $\mathbf{w}^\nu$ and $\mathbf{x}$ then leads to

$$\begin{aligned}I_1 &= \int_0^\infty du \int_0^\infty du' \int_{-\infty}^\infty \frac{dv}{2\pi} \int_{-\infty}^\infty \frac{dv'}{2\pi} e^{i[(u+\gamma)v+(u'-1)v']-\hat{g}(1-\gamma^2)v^2/2} \left(1+\frac{\hat{g}v'^2}{N}\right)^{-N/2} \\ &\overset{(N\to\infty)}{=} \int_0^\infty du \int_0^\infty du' \int_{-\infty}^\infty \frac{dv}{2\pi} \int_{-\infty}^\infty \frac{dv'}{2\pi} e^{i[(u+\gamma)v+(u'-1)v']-\hat{g}(1-\gamma^2)v^2/2-\hat{g}v'^2/2}.\end{aligned} \tag{25}$$

Finally, the integrals over v and v' can be performed exactly, then those over u and u' can be evaluated using the complementary error function, yielding the final result

$$I_1 = \frac{1}{4}\mathrm{erfc}\left(\frac{\gamma}{\sqrt{2\hat{g}(1-\gamma^2)}}\right)\mathrm{erfc}\left(-\frac{1}{\sqrt{2\hat{g}}}\right). \tag{26}$$

In a similar manner, the integrals in (23) can be evaluated to get

$$
\begin{aligned}
I_2 &= \int_0^\infty du \int_0^\infty du' \int_{-\infty}^\infty \frac{dv}{2\pi} \int_{-\infty}^\infty \frac{dv'}{2\pi} e^{i[uv+(u'+1)v']} \int \frac{d\mathbf{x}}{(2\pi)^{N/2}} e^{-\mathbf{x}^2/2} \\
&\quad \times \int \frac{d\mathbf{w}^\nu}{(2\pi\hat{g}/N)^{N/2}} \exp\left(-\frac{N}{2\hat{g}}|\mathbf{w}^\nu|^2 + [\gamma v - v']\mathbf{w}^\nu \cdot \mathbf{x} - \frac{\hat{g}(1-\gamma^2)}{2} v^2\right) \\
&= \int_0^\infty du \int_0^\infty du' \int_{-\infty}^\infty \frac{dv}{2\pi} \int_{-\infty}^\infty \frac{dv'}{2\pi} e^{i[uv+(u'+1)v']} \left[1 + \frac{\hat{g}}{N}(v^2 + v'^2 - 2\gamma v v')\right]^{-N/2} \\
&\overset{(N\to\infty)}{=} \int_0^\infty du \int_0^\infty du' \int_{-\infty}^\infty \frac{dv}{2\pi} \int_{-\infty}^\infty \frac{dv'}{2\pi} \exp\left(i[uv + (u'+1)v'] - \frac{\hat{g}}{2}(v^2 + v'^2 - 2\gamma v v')\right).
\end{aligned}
\tag{27}
$$

After performing the integrals over v, v', and u', then changing variables for the u integral, the final result is

$$
I_2 = \frac{1}{\sqrt{8\pi}} \int_0^\infty dr\ e^{-r^2/2} \mathrm{erfc}\left(\frac{\gamma r + 1/\sqrt{\hat{g}}}{\sqrt{2(1-\gamma^2)}}\right), \tag{28}
$$

where the final integral in this case must be performed numerically.

As a check, we can evaluate these results in their extreme limits. In the case $\gamma = 1$, which corresponds to testing the most recently learned pattern, we have $I_1 \sim I_2 \sim \mathrm{erfc}(\infty) \to 0$, so that there is perfect classification for very recently learned patterns. In the opposite limit of $\gamma = 0$, which corresponds to testing patterns learned in the distant past, we obtain $I_1 + I_2 = \frac{1}{4}[\mathrm{erfc}(-1/\sqrt{\hat{g}}) + \mathrm{erfc}(1/\sqrt{\hat{g}})] = \frac{1}{2}$, which means that very old patterns are completely overwritten and so are classified at chance level.

Appendix 2: The two-pathway forgetting curve

Let us introduce a second source of input to the downstream units, so that $z^\mu = \phi(\mathbf{w} \cdot \mathbf{x}^\mu + \mathbf{v} \cdot \mathbf{y}^\mu)$, where $x_i^\mu, y_i^\mu \sim \mathcal{N}(0,1)$. Though it is not necessary, we will assume for notational simplicity that the numbers of units in the two input layers are the same, so that $N_x = N_y = N$. The weights $\mathbf{w}$ are again trained using supervised learning:

$$
\Delta\mathbf{w}^\mu = \begin{cases} (\kappa\hat{z}^\mu - \mathbf{w}^\mu \cdot \mathbf{x}^\mu - \mathbf{v}^\mu \cdot \mathbf{y}^\mu)\mathbf{x}^\mu/N, & (\mathbf{w}^\mu \cdot \mathbf{x}^\mu + \mathbf{v}^\mu \cdot \mathbf{y}^\mu)\hat{z}^\mu < \kappa, \\ 0, & \text{else}, \end{cases} \tag{29}
$$

The second set of weights, meanwhile, is updated using the following rule:

$$
\Delta\mathbf{v}^\mu = -\frac{\alpha n_\mu}{N\bar{n}}\mathbf{v}^\mu + \sqrt{2}\frac{\beta n_\mu}{N\bar{n}}\hat{z}^\mu \mathbf{y}^\mu. \tag{30}
$$

This Hebbian update rule defines an N-dimensional Ornstein-Uhlenbeck stochastic process with time-dependent coefficients. The evolution of the probability distribution $p(\mathbf{v}, \Delta t)$ is given by the master equation:

$$
p(\mathbf{v}, t + \Delta t) = \int d\Delta\mathbf{v} p(\Delta\mathbf{v}|\mathbf{v} - \Delta\mathbf{v}) p(\mathbf{v} - \Delta\mathbf{v}, t), \tag{31}
$$

where

$$
p(\Delta\mathbf{v}|\mathbf{v} - \Delta\mathbf{v}) = \mathcal{N}\left(-\frac{\alpha n(t)}{N\bar{n}}(\mathbf{v} - \Delta\mathbf{v}), 2\left(\frac{\beta n(t)}{N\bar{n}}\right)^2\right). \tag{32}
$$

We can further note that, since the components of $\mathbf{v}$ are not coupled to one another in (30), we can, without loss of generality, consider the evolution of just a single component v_i. In this case the two sides of (31) can

be expanded to obtain

$$p(v_i,t)+\Delta t\frac{\partial p(v_i,t)}{\partial t}+O(\Delta t^2)=p(v_i,t)-\frac{\partial p(v_i,t)}{\partial v_i}\int d\Delta v_i p(\Delta v_i|v_i-\Delta v_i)\Delta v_i \\ +\frac{1}{2}\frac{\partial^2}{\partial v_i^2}p(v_i,t)\int d\Delta v_i p(\Delta v_i|v_i-\Delta v_i)(\Delta v_i)^2+O((\Delta v_i)^3). \tag{33}$$

Then, using (32) to obtain $\langle\Delta v_i\rangle$ and $\langle(\Delta v_i)^2\rangle$, (33) leads to the Fokker-Planck equation, which describes the evolution of the probability distribution $p(\mathbf{v})$ as new patterns are learned:

$$\frac{dp(\mathbf{v})}{dt}=\frac{\alpha n(t)}{\bar{n}}\sum_i\frac{\partial}{\partial v_i}[v_i p(\mathbf{v})]+\frac{(\beta n(t)/\bar{n})^2}{N}\sum_i\frac{\partial^2 p(\mathbf{v})}{\partial v_i^2}. \tag{34}$$

In this equation, we have taken the continuous-time limit by letting $\Delta t=1/N$ and $t\equiv(\mu-\nu)/N$, where $\mu>\nu$, the number of repetitions to be $n(t)\equiv n_\mu$, and the initial condition to be given by the distribution (to be calculated below) $p(\mathbf{v}^\nu)$.

The Fokker-Planck equation (34) can be solved using the Fourier transform [54]

$$p(\mathbf{k})=\int d\mathbf{v}e^{-i\mathbf{k}\cdot\mathbf{v}}p(\mathbf{v}). \tag{35}$$

With this, (34) becomes

$$\frac{dp(\mathbf{k})}{dt}=-\frac{\alpha n(t)}{\bar{n}}\mathbf{k}\cdot\nabla_{\mathbf{k}}p(\mathbf{k})-\frac{(\beta n(t)/\bar{n})^2}{N}|\mathbf{k}|^2p(\mathbf{k}). \tag{36}$$

This equation can be solved by making the following ansatz:

$$p(\mathbf{k})=\exp\left(-i\mathbf{k}\cdot\mathbf{m}(t)-\frac{1}{2}\mathbf{k}\cdot\mathbf{S}(t)\cdot\mathbf{k}\right), \tag{37}$$

which, by substituting into (36) and requiring that the terms at each order in $\mathbf{k}$ vanish, leads to the following equations for the time-dependent coefficients:

$$\begin{aligned}\frac{dm_i}{dt}&=-\frac{\alpha n(t)}{\bar{n}}m_i\\ \frac{dS_{ij}}{dt}&=-2\frac{\alpha n(t)}{\bar{n}}S_{ij}+2\frac{(\beta n(t)/\bar{n})^2}{N}\delta_{ij}.\end{aligned} \tag{38}$$

These equations then have the solutions

$$\begin{aligned}m_i(t)&=m_i(0)\exp\left(-\frac{\alpha}{\bar{n}}\int_0^t dt' n(t')\right)\\ S_{ij}(t)&=2\delta_{ij}\frac{\beta^2}{N\bar{n}^2}\int_0^t dt'\exp\left(-2\frac{\alpha}{\bar{n}}\int_{t'}^t dt'' n(t'')\right)n^2(t-t')\end{aligned} \tag{39}$$

With this, and letting $t=(P-\mu)/N$, we can take the inverse Fourier transform of (37) to obtain the distribution of the final weight vector given the weights at pattern ν:

$$p(\mathbf{v}^P|\mathbf{v}^\nu)=\frac{1}{(2\pi\sigma_P^2/N)^{N/2}}\exp\left(-\frac{N}{2\sigma_P^2}[\mathbf{v}^P-\rho\mathbf{v}^\nu]^2\right), \tag{40}$$

where we have identified $S_{ij}(t)=\sigma_P^2(P,\nu)\delta_{ij}$, with

$$\sigma_P^2(P,\nu)\equiv\frac{2\beta^2}{\bar{n}^2N}\sum_{\mu=\nu}^P\rho^2(P,\mu)n_\mu^2, \tag{41}$$

and we have also identified $m_i(t) = v_i^\nu \rho(P,\nu)$, with

$$\rho(P,\nu) \equiv \exp\left(-\frac{\alpha}{N\bar{n}}\sum_{\mu=\nu}^{P} n_\mu\right). \tag{42}$$

In what follows below, in order to keep expressions compact, we shall write $\sigma_P^2 = \sigma_P^2(P,\nu)$ and $\rho = \rho(P,\nu)$. Further, though it is not strictly necessary, these expressions can be considerably simplified if we make the simplifying assumption that the average of n_μ in (42) over the last $P-\nu$ patterns is equal to its average $\bar{n}$ over the full set of P patterns (or, in the case of (41), that n_μ^2 can be replaced by $\bar{n}^2$). In this case, (42) becomes $\rho(P,\nu) = e^{-\alpha(P-\nu)/N}$, while (41), after performing the summation, becomes $\sigma_P^2 = \beta^2(1-\rho^2)/\alpha$.

Returning to the distribution (40), we can see that it begins as a δ function at $\mathbf{v}^\nu$ and, as more patterns are introduced and $\rho \to 0$, evolves to the following steady-state distribution:

$$p(\mathbf{v}) = \frac{1}{(2\pi\beta^2/\alpha N)^{N/2}} e^{-\alpha N \mathbf{v}^2/2\beta^2}. \tag{43}$$

This is the distribution from which $\mathbf{v}^\nu$ will be drawn in order to calculate the error rate for pattern ν below. In addition to driving a drift-diffusion process for the weights $\mathbf{v}$, the updates to $\mathbf{v}$ also affect the evolution of the distribution of weights $\mathbf{w}$ due to the appearance of $\mathbf{v}$ in the update rule (29). In order to account for this change, the first two moments of the update $\Delta\mathbf{w}$ must be reevaluated as in (7)-(8), but now averaging over the random variable $\mathbf{y}$ in addition to averaging over $\mathbf{x}$. Because $\langle y_i \rangle = 0$, the first moment in (7) remains unchanged. As for the second moment, (8) generalizes to

$$\begin{aligned}
\Sigma_{ij}(\mathbf{w},\mathbf{v}) &\equiv \langle \Delta w_i \Delta w_j \rangle_{\mathbf{x},\mathbf{y}} - \mu_i(\mathbf{w})\mu_j(\mathbf{w}) \\
&= \frac{1}{N^2}\int d\mathbf{x}\; p(\mathbf{x}) \int d\mathbf{y}\; p(\mathbf{y})(1-\mathbf{w}\cdot\mathbf{x}-\mathbf{v}\cdot\mathbf{y})^2 x_i x_j - \mu_i(\mathbf{w})\mu_j(\mathbf{w}) \\
&= \frac{2}{N^2}\left(\left[g(|\mathbf{w}|) + \frac{|\mathbf{v}|^2}{2}\right]\delta_{ij} + h(|\mathbf{w}|) w_i w_j\right).
\end{aligned} \tag{44}$$

As before, we will assume that $|\mathbf{w}|^2$ and $|\mathbf{v}|^2$ can be replaced by their average values in the $N \to \infty$ limit. Noting from (43) that $\langle \mathbf{v}^2 \rangle = \beta^2/\alpha$, we have the diffusion tensor $\Sigma_{ij} = 2[\tilde{g}\delta_{ij} + \hat{h} w_i w_j]/N^2$, where $\tilde{g} \equiv g(\hat{w}) + \beta^2/2\alpha$, where, as before, $\hat{w} \equiv \sqrt{\langle |\mathbf{w}|^2 \rangle}$ is taken from numerical simulations. From this result, we see that the equations (17)-(18) determining the evolution of $\mathbf{w}$ can also be applied in the two-pathway case by making the substitution $g \to \tilde{g}$.

As in the previous section, our goal is to calculate the probability of incorrect classification when testing pattern ν after training P patterns in sequence. Including the second input pathway, this quantity is given by

$$\begin{aligned}
p(z^\nu \neq \hat{z}^\nu) = &\int d\mathbf{x}^\nu p(\mathbf{x}^\nu) \int d\mathbf{y}^\nu p(\mathbf{y}^\nu) \int d\mathbf{w}^\nu p(\mathbf{w}^\nu) \int d\mathbf{v}^\nu p(\mathbf{v}^\nu) \int d\mathbf{v}^P p\left(\mathbf{v}^P \middle| \mathbf{v}^\nu - \frac{\alpha n_\nu}{N\bar{n}}\mathbf{v}^\nu + \sqrt{2}\frac{\beta n_\nu}{N\bar{n}}\mathbf{y}^\nu\right) \\
&\times \int d\mathbf{w}^P \Theta(-\mathbf{w}^P\cdot\mathbf{x}^\nu - \mathbf{v}^P\cdot\mathbf{y}^\nu) \\
&\times \big[\Theta(1-\mathbf{w}^\nu\cdot\mathbf{x}^\nu - \mathbf{v}^\nu\cdot\mathbf{y}^\nu) p(\mathbf{w}^P|\mathbf{w}^\nu + [1-\mathbf{w}^\nu\cdot\mathbf{x}^\nu - \mathbf{v}^\nu\cdot\mathbf{y}^\nu]\mathbf{x}^\nu/N) \\
&\qquad + \Theta(\mathbf{w}^\nu\cdot\mathbf{x}^\nu + \mathbf{v}^\nu\cdot\mathbf{x}^\nu - 1) p(\mathbf{w}^P|\mathbf{w}^\nu)\big].
\end{aligned} \tag{45}$$

As before, we have absorbed all $\hat{z}^\mu$ into the definition of the input activity vectors $\mathbf{x}^\mu$ and $\mathbf{y}^\mu$ by letting $\mathbf{x}^\mu \to \hat{z}^\mu \mathbf{x}^\mu$ and $\mathbf{y}^\mu \to \hat{z}^\mu \mathbf{y}^\mu$, effectively setting all $\hat{z}^\mu = 1$. The first term in the integrand corresponds to cases in which the classification of pattern ν is initially incorrect, so that the weights $\mathbf{w}$ and $\mathbf{v}$ are both updated. The second term corresponds to cases in which the initial classification is correct, so that only the

weights $\mathbf{v}$ are updated. In both cases, the weight distributions for $\mathbf{w}$ and $\mathbf{v}$ evolve according to drift-diffusion processes, as new patterns are learned up until pattern P, at which time, the classification of pattern ν is tested using the final weights $\mathbf{w}^P$ and $\mathbf{v}^P$ (the first Θ function appearing in the integrand).

Again using the trick (20) to represent the Heaviside step functions, the two terms in (45) can be written as

$$p(z^\nu \neq \hat{z}^\nu) = J_1 + J_2, \tag{46}$$

where

$$\begin{aligned} J_1 = &\int_0^\infty du \int_0^\infty du' \int \frac{dv}{2\pi} \int \frac{dv'}{2\pi} e^{i[vu+v'(u'-1)]} \int d\mathbf{x}\; p(\mathbf{x}) \int d\mathbf{y}\; p(\mathbf{y}) \\ &\times \int d\mathbf{w}^\nu p^*(\mathbf{w}^\nu) \int d\mathbf{v}^\nu p^*(\mathbf{v}^\nu) \int d\mathbf{v}^P p\left(\mathbf{v}^P \middle| \mathbf{v}^\nu - \frac{\alpha n_\nu}{N\bar{n}}\mathbf{v}^\nu + \sqrt{2}\frac{\beta n_\nu}{N\bar{n}}\mathbf{y}\right) \\ &\times \int d\mathbf{w}^P p(\mathbf{w}^P|\mathbf{w}^\nu + [1 - \mathbf{w}^\nu\cdot\mathbf{x}^\nu - \mathbf{v}^\nu\cdot\mathbf{y}^\nu]\mathbf{x}^\nu/N) e^{i[(v\mathbf{w}^P + v'\mathbf{w}^\nu)\cdot\mathbf{x} + (v\mathbf{v}^P + v'\mathbf{v}^\nu)\cdot\mathbf{y}]} \end{aligned} \tag{47}$$

and

$$\begin{aligned} J_2 = &\int_0^\infty du \int_0^\infty du' \int \frac{dv}{2\pi} \int \frac{dv'}{2\pi} e^{i[vu+v'(u'+1)]} \int d\mathbf{x}\; p(\mathbf{x}) \int d\mathbf{y}\; p(\mathbf{y}) \\ &\times \int d\mathbf{w}^\nu p^*(\mathbf{w}^\nu) \int d\mathbf{v}^\nu p^*(\mathbf{v}^\nu) \int d\mathbf{v}^P p\left(\mathbf{v}^P \middle| \mathbf{v}^\nu - \frac{\alpha n_\nu}{N\bar{n}}\mathbf{v}^\nu + \sqrt{2}\frac{\beta n_\nu}{N\bar{n}}\mathbf{y}\right) \\ &\times \int d\mathbf{w}^P p(\mathbf{w}^P|\mathbf{w}^\nu) e^{i[(v\mathbf{w}^P - v'\mathbf{w}^\nu)\cdot\mathbf{x} + (v\mathbf{v}^P - v'\mathbf{v}^\nu)\cdot\mathbf{y}]}. \end{aligned} \tag{48}$$

Beginning with J_1, we can shift the integration variable $\mathbf{w}^P \to \mathbf{w}^P - \gamma(\mathbf{v}^\nu\cdot\mathbf{y})\mathbf{x}/N$ and use $\mathbf{x}^2 = N$ in the $N \to \infty$ limit to obtain

$$\begin{aligned} J_1 \overset{(N\to\infty)}{=} &\int_0^\infty du \int_0^\infty du' \int \frac{dv}{2\pi} \int \frac{dv'}{2\pi} e^{i[vu+v'(u'-1)]} \int d\mathbf{x}\; p(\mathbf{x}) \int d\mathbf{w}^\nu p^*(\mathbf{w}^\nu) \\ &\times \int d\mathbf{w}^P p(\mathbf{w}^P|\mathbf{w}^\nu + [1 - \mathbf{w}^\nu\cdot\mathbf{x}]\mathbf{x}/N) e^{i(v\mathbf{w}^P + v'\mathbf{w}^\nu)\cdot\mathbf{x}} K_1(v, v'), \end{aligned} \tag{49}$$

where

$$\begin{aligned} K_1(v,v') \equiv &\int d\mathbf{y}\; p(\mathbf{y}) \int d\mathbf{v}^\nu p^*(\mathbf{v}^\nu) \int d\mathbf{v}^P p\left(\mathbf{v}^P \middle| \mathbf{v}^\nu + \sqrt{2}\frac{\beta n_\nu}{N\bar{n}}\mathbf{y}\right) e^{i[v\mathbf{v}^P + (v'-\gamma v)\mathbf{v}^\nu]\cdot\mathbf{y}} \\ = &\int \frac{d\mathbf{y}}{(2\pi)^{N/2}} e^{-\mathbf{y}^2/2} \int \frac{d\mathbf{v}^\nu}{(2\pi\beta^2/\alpha N)^{N/2}} e^{-\alpha N|\mathbf{v}^\nu|^2/2\beta^2} \\ &\times \int \frac{d\mathbf{v}^P}{(2\pi\sigma_P^2/N)^{N/2}} \exp\left(-\frac{N}{2\sigma_P^2}\left[\mathbf{v}^P - \rho\left(\mathbf{v}^\nu + \sqrt{2}\frac{\beta n_\nu}{N\bar{n}}\mathbf{y}\right)\right]^2\right) e^{i[v\mathbf{v}^P + (v'-\gamma v)\mathbf{v}^\nu]\cdot\mathbf{y}} \\ = &\int \frac{d\mathbf{y}}{(2\pi)^{N/2}} e^{-\mathbf{y}^2/2} \int \frac{d\mathbf{v}^\nu}{(2\pi\beta^2/\alpha N)^{N/2}} \int \frac{d\mathbf{v}^P}{(2\pi\sigma_P^2/N)^{N/2}} \\ &\times \exp\left(-\frac{\alpha N}{2\beta^2}|\mathbf{v}^\nu|^2 - \frac{N}{2\sigma_P^2}|\mathbf{v}^P|^2 + i\mathbf{y}\cdot\left[v\mathbf{v}^P + (v' + (\rho-\gamma)v)\mathbf{v}^\nu + \sqrt{2}\frac{\rho\beta n_\nu}{N\bar{n}}v\mathbf{y}\right]\right) \\ = &\int \frac{d\mathbf{y}}{(2\pi)^{N/2}} \exp\left(-\frac{\mathbf{y}^2}{2}\left[1 + \frac{1}{N}\left(\sigma_P^2 v^2 + \frac{\beta^2}{\alpha}(v' + (\rho-\gamma)v)^2 - 2\sqrt{2}i\frac{\rho\beta n_\nu}{\bar{n}}v\right)\right]\right) \\ = &\left[1 + \frac{1}{N}\left(\sigma_P^2 v^2 + \frac{\beta^2}{\alpha}(v' + (\rho-\gamma)v)^2 - 2\sqrt{2}i\frac{\rho\beta n_\nu}{\bar{n}}v\right)\right]^{-N/2} \\ \overset{(N\to\infty)}{=} &\exp\left(-\frac{\sigma_P^2}{2}v^2 - \frac{\beta^2}{2\alpha}[v' + (\rho-\gamma)v)]^2 + \sqrt{2}i\frac{\beta\rho n_\nu}{\bar{n}}v\right) \end{aligned} \tag{50}$$

Where the Hebbian decay term $\sim \alpha \mathbf{v}^\nu/N$ was dropped in the first line because it vanishes as $N \to \infty$, and the integration variable $\mathbf{v}^P$ was shifted in the third line. Using this result, and noting that the integrals over $\mathbf{x}$, $\mathbf{w}^\nu$, and $\mathbf{w}^P$ are the same as those appearing in (22) (with $\hat{g} \to \tilde{g}$), (49) becomes

$$\begin{aligned} J_1 &= \int_0^\infty du \int_0^\infty du' \int \frac{dv}{2\pi} \int \frac{dv'}{2\pi} \exp\Bigg(-\frac{v^2}{2}\left[(1-\gamma^2)\tilde{g} + \sigma_P^2 + (\rho-\gamma)^2\frac{\beta^2}{\alpha}\right] - \frac{v'^2}{2}\left[\tilde{g} + \frac{\beta^2}{\alpha}\right] \\ &\qquad + iv\left[u + \gamma + \sqrt{2}\frac{\beta\rho n_\nu}{\bar{n}} + i(\rho-\gamma)\frac{\beta^2}{\alpha}v'\right] + iv'[u'-1]\Bigg) \\ &= \frac{1}{\sqrt{8\pi}}\int_{r_1}^\infty dr\ e^{-r^2/2}\mathrm{erfc}\left(\frac{s_1(r)}{\sqrt{2}}\right), \end{aligned} \tag{51}$$

where we have defined

$$r_1 \equiv \frac{\gamma + \sqrt{2}\beta\rho n_\nu/\bar{n}}{\sqrt{(1-\gamma^2)\tilde{g} + \sigma_P^2 + (\rho-\gamma)^2\beta^2/\alpha}} \tag{52}$$

and

$$s_1(r) = \frac{-\sqrt{(1-\gamma^2)\tilde{g} + \sigma_P^2 + (\rho-\gamma)^2\beta^2/\alpha} + r(\gamma-\rho)\beta^2/\alpha}{\sqrt{\left(\tilde{g} + \frac{\beta^2}{\alpha}\right)\left[(1-\gamma^2)\tilde{g} + \sigma_P^2 + (\rho-\gamma)^2\beta^2/\alpha\right] - (\rho-\gamma)^2\frac{\beta^4}{\alpha^2}}}. \tag{53}$$

As in the case without the Hebbian pathway, the last remaining integral in (51) must be performed numerically.

Equation (48) for J_2 can be evaluated in a similar manner. To begin, we express it as

$$\begin{aligned} J_2 &\overset{(N\to\infty)}{=} \int_0^\infty du \int_0^\infty du' \int \frac{dv}{2\pi} \int \frac{dv'}{2\pi} e^{i[vu + v'(u'+1)]} \int d\mathbf{x}\ p(\mathbf{x}) \int d\mathbf{w}^\nu p^*(\mathbf{w}^\nu) \\ &\quad \times \int d\mathbf{w}^P p\left(\mathbf{w}^P|\mathbf{w}^\nu\right) e^{i(v\mathbf{w}^P - v'\mathbf{w}^\nu)\cdot\mathbf{x}} K_2(v,v'), \end{aligned} \tag{54}$$

where

$$\begin{aligned} K_2(v,v') &\equiv \int d\mathbf{y}\ p(\mathbf{y}) \int d\mathbf{v}^\nu p^*(\mathbf{v}^\nu) \int d\mathbf{v}^P p\left(\mathbf{v}^P \middle| \mathbf{v}^\nu + \sqrt{2}\frac{\beta n_\nu}{N\bar{n}}\mathbf{y}\right) e^{i(v\mathbf{v}^P - v'\mathbf{v}^\nu)\cdot\mathbf{y}} \\ &= \int \frac{d\mathbf{y}}{(2\pi)^{N/2}} e^{-\mathbf{y}^2/2} \int \frac{d\mathbf{v}^\nu}{(2\pi\beta^2/N\alpha)^{N/2}} e^{-\alpha N|\mathbf{v}^\nu|^2/2\beta^2} \\ &\quad \times \int \frac{d\mathbf{v}^P}{(2\pi\sigma_P^2/N)^{N/2}} \exp\left(-\frac{N}{2\sigma_P^2}\left[\mathbf{v}^P - \rho\left(\mathbf{v}^\nu + \sqrt{2}\frac{\beta n_\nu}{N\bar{n}}\mathbf{y}\right)\right]^2\right) e^{i(v\mathbf{v}^P - v'\mathbf{v}^\nu)\cdot\mathbf{y}} \\ &= \int \frac{d\mathbf{y}}{(2\pi)^{N/2}} \exp\left(-\frac{\mathbf{y}^2}{2}\left[1 + \frac{1}{N}\left(\sigma_P^2 v^2 + \frac{\beta^2}{\alpha}(\rho v - v')^2 - 2\sqrt{2}i\frac{\rho\beta n_\nu}{\bar{n}}v\right)\right]\right) \\ &= \left[1 + \frac{1}{N}\left(\sigma_P^2 v^2 + \frac{\beta^2}{\alpha}(\rho v - v')^2 - 2\sqrt{2}i\frac{\rho\beta n_\nu}{\bar{n}}v\right)\right]^{-N/2} \\ &\overset{(N\to\infty)}{=} \exp\left(-\frac{\sigma_P^2}{2}v^2 - \frac{\beta^2}{2\alpha}(\rho v - v')^2 + \sqrt{2}i\frac{\rho\beta n_\nu}{\bar{n}}v\right) \end{aligned} \tag{55}$$

Using this result, and noting that the integrals over $\mathbf{x}$, $\mathbf{w}^\nu$, and $\mathbf{w}^P$ are the same as those appearing in (23) (with $\hat{g} \to \tilde{g}$), (54) becomes

$$\begin{aligned} J_2 &= \int_0^\infty du \int_0^\infty du' \int \frac{dv}{2\pi} \int \frac{dv'}{2\pi} \exp\left(i[uv + (u'+1)v'] - \frac{\tilde{g}}{2}[v^2 + v'^2 - 2\gamma vv']\right) \\ &\quad \times \exp\left(-\frac{\sigma_P^2}{2}v^2 - \frac{\beta^2}{2\alpha}(\rho v - v')^2 + \sqrt{2}i\frac{\rho\beta n_\nu}{\bar{n}}v\right) \\ &= \frac{1}{\sqrt{8\pi}}\int_{r_2}^\infty dr\ e^{-r^2/2}\mathrm{erfc}\left(\frac{s_2(r)}{\sqrt{2}}\right), \end{aligned} \tag{56}$$

where we have defined

$$r_2 = \frac{\sqrt{2}\rho\beta n_\nu/\bar{n}}{\sqrt{\tilde{g} + \sigma_P^2 + \rho^2\beta^2/\alpha}} \tag{57}$$

and

$$s_2(r) = \frac{\sqrt{\tilde{g} + \sigma_P^2 + \rho^2\beta^2/\alpha} + (\gamma\tilde{g} + \rho\beta^2/\alpha)r}{\sqrt{\left(\tilde{g} + \frac{\beta^2}{\alpha}\right)(\tilde{g} + \sigma_P^2 + \rho^2\beta^2/\alpha) - (\gamma\tilde{g} + \rho\beta^2/\alpha)^2}}. \tag{58}$$

It is straightforward to check that, in the limit $\beta \to 0$, which corresponds to shutting off the Hebbian input pathway, we recover $J_{1,2} \to I_{1,2}$ from the simple perceptron result.

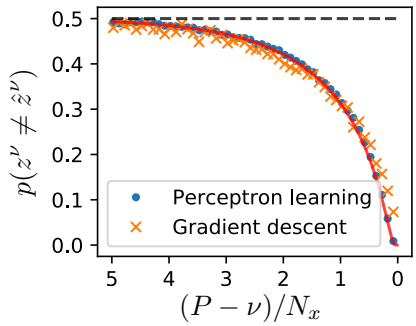
Finally, we can calculate the probability of making a weight update in a given step, which in the two-pathway case is given by

$$\begin{aligned} q &= \int \frac{d\mathbf{x}}{(2\pi)^{N/2}} e^{-|\mathbf{x}|^2/2} \int \frac{d\mathbf{y}}{(2\pi)^{N/2}} e^{-|\mathbf{y}|^2/2} \int \frac{d\mathbf{w}}{(2\pi\hat{w}^2/N)^{N/2}} e^{-N|\mathbf{w}|^2/2\hat{w}^2} \\ &\quad \times \int \frac{d\mathbf{v}}{(2\pi\beta^2/\alpha N)^{N/2}} e^{-N\alpha|\mathbf{v}|^2/2\beta^2} \Theta(1 - \mathbf{w}\cdot\mathbf{x} - \mathbf{v}\cdot\mathbf{y}) \\ &= \frac{1}{2}\mathrm{erfc}\left(-\frac{1}{\sqrt{2(\hat{w}^2 + \beta^2/\alpha)}}\right). \end{aligned} \tag{59}$$

In the case where $\beta^2/\alpha = 0$, this evaluates to the earlier result $q \simeq 0.798$. For nonzero β, on the other hand, q decreases and approaches 0.5 as $\beta^2/\alpha \to \infty$.

Supplemental Information

Learning with gradient descent



Supplemental Figure 1: Training with gradient descent yields a forgetting curve similar to the perceptron case.

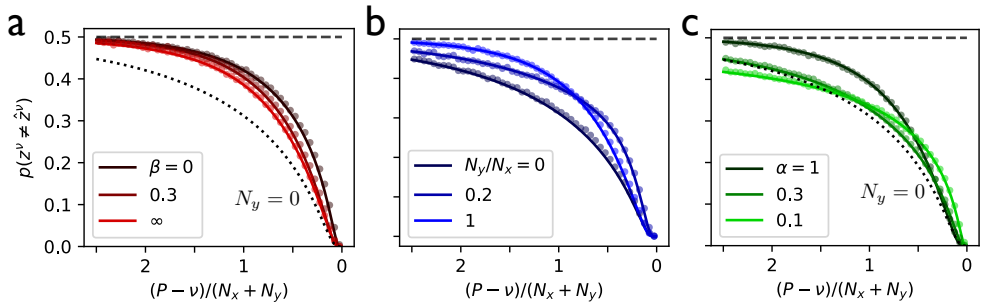
The perceptron learning rule, according to which synaptic weights are adjusted to produce the correct output in a single step, is mathematically convenient but biologically questionable. In order to address this, we simulated the forgetting curve using gradient descent learning, a widely used supervised learning algorithm in which small updates are accumulated over many repetitions to minimize the readout error.

In this case the output of the neuron is $z_a^\mu = \mathrm{sgn}(\mathbf{w}^\mu \cdot \mathbf{x}^\mu + \sigma\xi_a)$, where $\xi_a \sim \mathcal{N}(0,1)$ is drawn randomly for step a. The number of steps for each pattern was chosen to be $N_{\mathrm{steps}} = 100$, and the learning rate $\eta = 0.01$ and the noise amplitude $\sigma = 0.2$ were chosen by grid search to maximize the area between the forgetting curve and chance performance. The noise is not strictly necessary in gradient descent learning, but is included here so that a finite classification margin will be obtained, as in the case of the perceptron learning rule (Figure 1b). In this case, the synaptic weights were updated according to the gradient-descent update rule $\Delta w_i^\mu = \sum_{a=1}^{N_{\mathrm{steps}}} \delta w_{i,a}^\mu$, where

$$\delta w_{i,a}^\mu = \frac{\eta}{N_x}[\hat{z}^\mu - z_a^\mu]x_i^\mu. \tag{60}$$

As shown in Supplemental Figure 1, the forgetting curve obtained for the perceptron with this alternative learning rule is similar to that obtained using the perceptron learning rule.

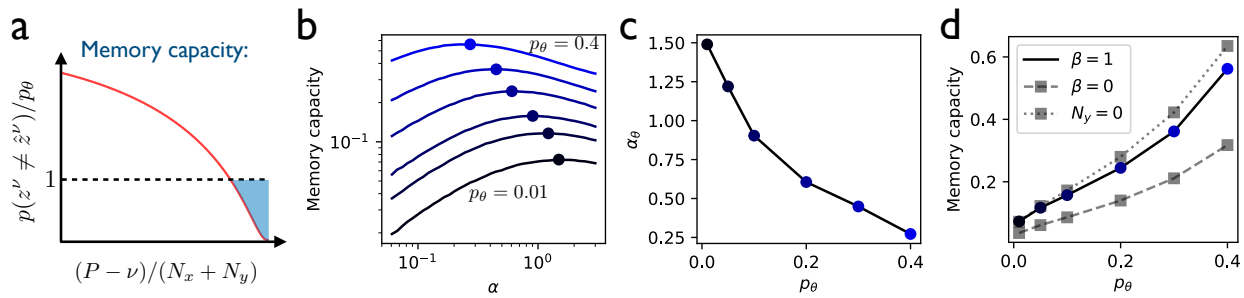
Two-pathway forgetting curves depend on Hebbian learning and decay rates



Supplemental Figure 2: Forgetting curves in the two-pathway model with each pattern trained once. **(a)** The forgetting curve for different values of the Hebbian learning rate β, with $\alpha = 1$ and $N_x = N_y$ (dotted line shows the case with no second pathway). **(b)** The forgetting curve for different values of N_y/N_x, with $\alpha = \beta = 1$. **(c)** The forgetting curve for different values of the Hebbian weight decay rate α, with $\beta = 1$ and $N_x = N_y$. In (b)-(d), solid lines are theoretical results; points are simulations with $N_x = N_y = 1000$.

In the two-pathway model, we kept $n_\nu = \bar{n}$ constant for all patterns and investigated the dependence of the forgetting curve (4) on its other parameters. Starting with the Hebbian learning rate, we found that nonzero values of β shifted the forgetting curve slightly downward, modestly reducing the error rate for all patterns (Figure 2b). Whether this qualifies as a true improvement, however, depends somewhat on bookkeeping. For a fixed total number of synapses $N_x + N_y$, the error rate can be reduced by allowing for Hebbian learning. However, the error rate is reduced even more by eliminating the Hebbian synapses entirely (dashed curve in Figure 2b), which decreases the denominator $N_x + N_y$ by setting $N_y = 0$. Stated differently, if the goal is to minimize the error rate for a fixed total number of synapses, this is accomplished most effectively by letting all of the synapses be updated with supervised learning rather than with Hebbian learning (Figure 2c). If, on the other hand, the goal is to minimize the error rate for a fixed number N_x of supervised synapses, then a benefit is obtained from including additional synapses with Hebbian learning. The exception to these conclusions occurs for small values of the Hebbian decay rate α, in which case very old memories can persist for longer with error rates below chance level (Figure 2d). In this case, there is a benefit to adding Hebbian synapses even if they are counted against the total $N_x + N_y$.

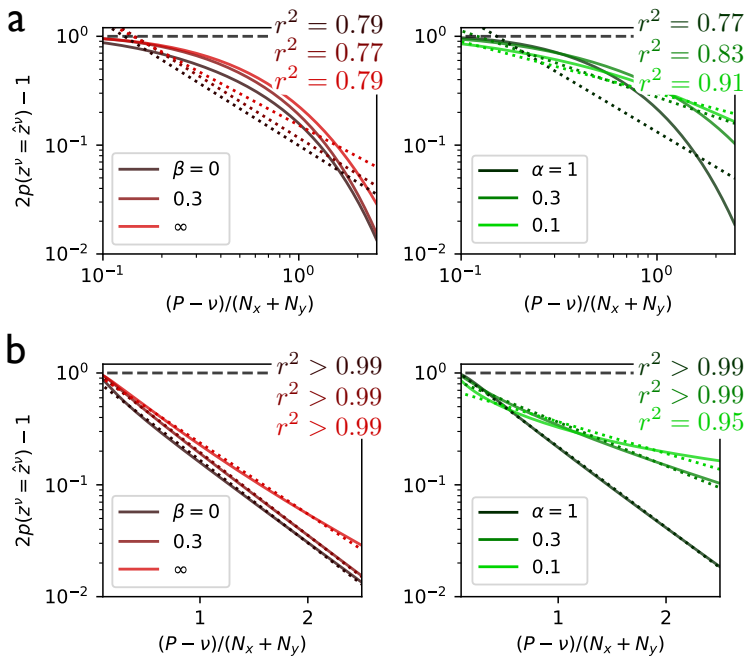
Optimal forgetting rate and memory capacity



Supplemental Figure 3: Error tolerance determines the optimal forgetting rate and memory capacity. **(a)** Given a threshold p_θ of incorrect classification probability, the memory capacity is defined as the area above the forgetting curve and below the threshold, normalized by p_θ. **(b)** For each value of p_θ, the memory capacity is optimized with respect to the Hebbian forgetting rate α. (For all curves, $\beta = 1$ is fixed, $N_x = N_y$, and $n_\nu = \bar{n}$ for all patterns.) **(c)** The optimal forgetting rate from (b) as a function of p_θ. **(d)** The memory capacity from (b) as a function of the classification threshold p_θ. Dashed line shows the capacity with a second pathway with no learning ($N_y = N_x$, $\beta = 0$); Dotted line shows the capacity with no second pathway ($N_y = 0$).

We investigated the effects of the parameters α and β in the two-pathway model, while still holding n^ν constant for all patterns, by setting a threshold p_θ for the acceptable error rate, then defining the memory capacity as the integrated area between the forgetting curve and this threshold value, normalized by p_θ (Supplemental Figure 3a). Because the forgetting curve was found to have relatively weak dependence on β (Figure 2b), we fixed $\beta = 1$ and considered the effects of α and of p_θ on the memory capacity. For a given choice of p_θ, we found the value of α that maximized the memory capacity (Supplemental Figure 3b). This led to the conclusion that, the larger the error tolerance p_θ, the smaller the forgetting rate α should be in order to maximize the memory capacity (Supplemental Figure 3c). Finally, we found that the optimized memory capacity increases as the error tolerance becomes greater (Supplemental Figure 3d). Consistent with our observations from Figure 2, we found that the memory capacity, which is normalized by the total number of synapses $N_x + N_y$, is improved by learning in the second pathway if $N_x + N_y$ is held constant (solid vs. dashed curve in Supplemental Figure 3d), but is even larger if the second pathway is left out entirely (dotted curve in Supplemental Figure 3d), again indicating that adding supervised synapses is a better strategy than adding unsupervised synapses if the goal is to optimize the forgetting curve for a fixed total number of synapses.

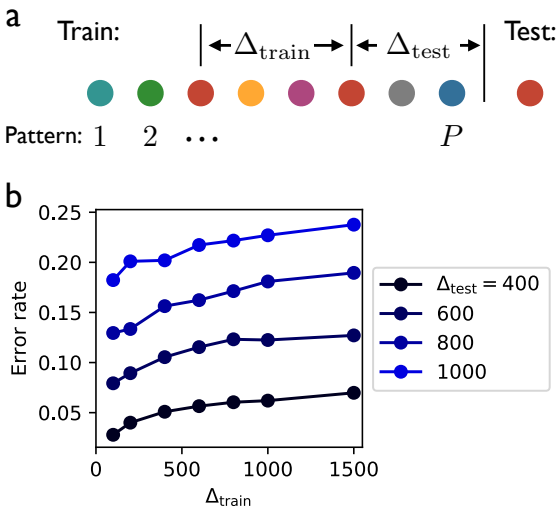
Forgetting curves decay exponentially



Supplemental Figure 4: Forgetting curves exhibit approximately exponential decay. **(a)** Forgetting curves from Figure 2b (*left*) and 2d (*right*) plotted on logarithmic axes, with $N_x = N$. Dotted lines are best fits with power law decay. **(b)** Forgetting curves as in (a), but plotted with just one logarithmic axis. Dotted lines are best fits with exponential decay.

Many previous studies on memory have shown that forgetting curves are well described by curves decaying as a power law [45], in which the probability of correct recall has the form $\sim t^{-a}$, where t is time and $a > 0$. Supplemental Figure 4a shows that such a fitting function, which appears as a straight line in a log-log plot, does a relatively poor job of fitting the forgetting curves from the two-pathway model. In contrast, exponentially decaying functions, in which the probability of correct recall $\sim e^{-at}$ and which appear as straight lines on semilogarithmic plots, provide a good fit to the forgetting curves (Supplemental Figure 4b).

Spaced repetition



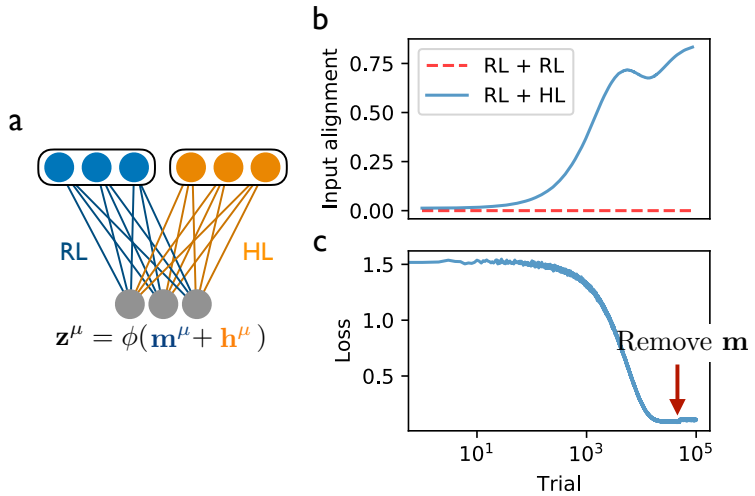
Supplemental Figure 5: Short training intervals during spaced repetition lead to better testing performance. **(a)** The two-pathway network is trained with random sequential patterns, with one particular pattern presented twice. **(b)** For any interval Δ_{test} between the second presentation and the testing phase, the testing performance for the repeated pattern is best when the interval between the presentations during training is short. Results are simulations from a two-pathway network with $N_x = N_y = 1000$ and $\alpha = \beta = 1$.

In simulations of the two-pathway model, the neuron was trained in sequence to perform P classifications. All of the input patterns were distinct, except for a single pattern that was presented twice during training, with an interval Δ_{train} between presentations. The interval between the second presentation and the testing phase was Δ_{test}. For any Δ_{test}, smaller values of Δ_{train} always led to a lower error rate for the repeated pattern during testing (Figure 5b).

In addition, we note that repetition of patterns has no significant effect at all in the single-pathway model without Hebbian weights. In this case, upon the second presentation of the repeated pattern during training, the weight vector will either not be updated at all (if classification is already correct with sufficient margin) or will be updated to lie on the correct side of the classification boundary plus a margin (if the classification is initially incorrect or correct with insufficient margin). In neither of these two cases is there a benefit to having seen the pattern before, since one of these same things would have happened if the first presentation had not occurred.

Thus, while the information accumulated in the Hebbian weights of the two-pathway model is an important ingredient for describing nontrivial effects due to spaced repetition, the single-neuron model appears to be unable to account for the experimentally observed existence of optimal repetition intervals [51], and thus leaves room for future work on this topic.

Input alignment with reinforcement learning



Supplemental Figure 6: Input alignment and control transfer in a network with reinforcement learning and Hebbian learning. **(a)** Version of the two-pathway model in which synaptic weights in the first pathway are updated with RL. **(b)** As training is repeated, the normalized alignment between $\mathbf{m}$ and $\mathbf{h}$ grows. The alignment does not occur if the second input pathway is updated with RL instead of Hebbian learning (dashed line) **(c)** The mean squared error between the output activity and the target state decreases with training and remains low if the first input pathway is removed.

We checked that the mechanisms of input alignment and control transfer could also be obtained using reinforcement learning (RL), which is a more plausible form of synaptic plasticity than the perceptron learning rule for describing changes at corticostriatal synapses. In this case, a noise current ξ was added to the readout layer, with $\xi_i \sim \mathcal{N}(0,1)$, and the updates to input weights in the first pathway were given by [55, 56]

$$\Delta W_{ij} = \eta(r - \bar{r})\xi_i. \tag{61}$$

Over the course of many trials, where each trial consists of a complete presentation of all $P = 10$ patterns, the input currents from the two pathways become aligned when one pathway was trained with RL and the other with HL, but not if both pathways were trained with RL (Supplemental Figure 6b). Further, due to input alignment and control transfer, the first input $\mathbf{m}$ can be removed once the network is trained, without significantly impairing performance (Supplemental Figure 6c).