HHS Public Access

Author manuscript

Behav Brain Res. Author manuscript; available in PMC 2021 December 31.

Published in final edited form as:

Behav Brain Res. 2019 December 30; 376: 112180. doi:10.1016/j.bbr.2019.112180.

Forgetting at biologically realistic levels of neurogenesis in a large-scale hippocampal model

Lina M. Tran[a,b,c], **Sheena A. Josselyn**[a,b,d,e,f], **Blake A. Richards**[g,i,j,k], **Paul W. Frankland**[a,b,d,e,h,*]

[a]Neuroscience and Mental Health, The Hospital for Sick Children, Toronto, ON, Canada

[b]Dept. of Physiology, University of Toronto, Toronto, ON, Canada

[c]Vector Institute, Toronto, ON, Canada

[d]Dept. of Psychology, University of Toronto, Toronto, ON, Canada

[e]Institute of Medical Sciences, University of TorontoToronto, ON, Canada

[f]Brain, Mind and Consciousness Program, Canadian Institute for Advanced Research, Toronto, ON, Canada

[g]Mila, Montreal, QC, Canada

[h]Child and Brain Development Program, Canadian Institute for Advanced Research, Toronto, ON, Canada

[i]School of Computer Science, McGill University, Montreal, QC, Canada

[j]Dept. of Neurology and Neurosurgery, McGill University, Montreal, QC, Canada

[k]Canadian Institute for Advanced Research, Toronto, ON, Canada

Abstract

Neurogenesis persists throughout life in the dentate gyrus region of the mammalian hippocampus. Computational models have established that the addition of neurons degrades existing memories (i.e., produces forgetting). These predictions are supported by empirical observations in rodents, where post-training increases in neurogenesis also promote forgetting of hippocampus-dependent memories. However, in these computational models which use 10–1,000 neurons to represent the dentate gyrus, forgetting is only observed at rates of new neuron addition that greatly exceed adult neurogenesis rates observed *in vivo*. In order to address this, here we generated an artificial neural network which incorporated more realistic features of the hippocampus – including increased network size (with up to 20,000 dentate gyrus neurons), sparse activity, and sparse connectivity – features that were not present in earlier models. In addition, we explored how properties of new neurons – their connectivity, excitability, and plasticity – impact forgetting using a pattern

*Corresponding author. paul.frankland@sickkids.ca (P.W. Frankland).

Declaration of interests

The authors declare no competing financial interests.

Appendix A. Supplementary Data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.bbr.2019.112180.

categorization task. Our results revealed that neurogenic networks forget previously learned input-output pattern associations. This forgetting predicted a performance enhancement in subsequent conflictual learning, compared to static networks (with no added neurons). These effects were especially sensitive to changes in increased output connectivity and excitability of new neurons. Crucially, forgetting was observed at much lower rates of neurogenesis in larger networks, with the addition of as little as 0.2% of the total DG population sufficient to induce forgetting.

Keywords

Neurogenesis; Forgetting; Artificial neural network

1. Introduction

The dentate gyrus (DG) of the hippocampus is distinct from many other regions in the brain, in so far as the production and integration of new neurons continues throughout life in humans [1–4] (but see [5]), and many non-human mammalian species [6–10]. Newly generated neurons are produced in the subgranular zone of the hippocampus, and within weeks they migrate and integrate into the DG. They transiently have distinct properties compared to developmentally-generated granule cells: they are more excitable (i.e., have a greater probability of firing) [11–13], more plastic (i.e., more readily undergo LTP) [14,15] and have different wiring properties than their mature counterparts [16,17]. The dentate gyrus has particularly sparse activity compared to other regions of the hippocampus [18]. Hence, a relatively active population of immature neurons in the DG is well-situated to impact activity locally within the DG as well as downstream. Accordingly, behavioral studies examining the impact of manipulating levels of hippocampal neurogenesis in adult rodents have suggested neurogenesis levels regulate a diverse collection of cognitive function in the hippocampus including pattern separation [19–21], forgetting [22], cognitive flexibility [23,24], modulation of systems consolidation [25] and mood regulation [26,27].

In order to understand how neurogenesis modulates hippocampal function, and memory in particular, various computational models have been explored [28–34]. These models have examined the impact of neurogenesis on new learning, as well as the stability of previously learned information. The general findings are that neurogenesis, on one hand, is beneficial for encoding new information and, on the other hand, may degrade stored memories (but see [31–33]). However, these previous models did not capture important features of the hippocampus, and, in particular, the DG: (1) The rodent DG contains > 1,000,000 granule cells [35]. Previous models used much lower numbers of neurons in the layer corresponding to the DG (10–1000); (2) In the rodent trisynaptic circuit, entorhinal cortex (EC) → DG → CA3, the approximate ratio of numbers of neurons in these subregions is 1:10:2 [36]. Previous models did not preserve the approximate ratio of neurons within the trisynaptic circuit. (3) Recurrent connections in CA3 allow for pattern completion. Previous models did not incorporate pattern completion capabilities in CA3. (4) Estimated rates of DG neurogenesis in adulthood are low, ranging from below 0.01% of the DG per day in macaques and humans [7,37,38] to 0.03–0.06% per DG per day in rodents. Previous models have observed effects on mnemonic function, but only after adding 5% to 30% of the DG

(across time or in one single epoch). Even the lower end of this range very much exceeds biological estimates.

Here we generated an artificial neural network (ANN) to explore the impact of neurogenesis on forgetting and also the interaction between forgetting and new learning. We established two main goals. First, to generate an ANN that incorporates the key features of the hippocampus, as outlined above, thereby permitting exploration of the impact of neurogenesis on hippocampal memory function under more physiologically-plausible conditions. Second, to use this model to systematically explore the impact of the unique properties of newborn neurons (connectivity, excitability, and plasticity) on memory function. Accordingly, our ANN was sparsely connected, had sparse activity, and layer sizes that mimicked EC→DG→CA3 neuroanatomy. Within our model we manipulated the excitability, plasticity, and connectivity of newborn neurons, and also varied network size (simulated DG size ranging from 500→20,000 neurons). The network was trained on a categorization task requiring pattern completion. Our main findings are that adding new neurons *after* original training promotes forgetting. These effects are observed regardless of network size, and are exacerbated by increased excitability and output connectivity of newborn neurons. In turn, neurogenesis-mediated forgetting facilitates new conflictual learning. Following forgetting of original categorizations, networks with neurogenesis exhibited facilitated reversal learning. Crucially, these same forgetting and facilitated reversal learning effects occur at low rates of neurogenesis (on par with biological estimates).

2. Methods

2.1. Three layer model of the hippocampus

We are using a feed forward artificial neural network (ANN) in which the input, hidden, and output layers are intended to model the EC, and DG and CA3 layers of the hippocampus. The network is a multilayer recurrent neural network, where the recurrent connections only exist in the output layer. This was implemented in order to observe pattern completion in the network in a manner that reflects the putative role of the CA3 in pattern completion in the hippocampus [39]. The network was built in Python 3.6 using Pytorch 0.4.1., and NumPy 1.14.3.

2.2. Model architecture

Based on relative numbers of neurons in the EC, DG and CA3 of rodents, the size of each layer, N, approximately reproduces the ratio of neuronal numbers between these regions (500, 5000, 1000 or 1:10:2).

Similarly, estimates from anatomical studies informed the choice of synaptic connectivity, *syn*, between these regions where *syn* is the proportion of target weights for each neuron in a layer to neurons in the next layer. For our network, we consider three values of *syn*: syn_i which is the proportion of target weights from the EC onto each DG neuron, syn_o which is the proportion of target weights from each DG neuron to the CA3, and syn_{CA3} which is the proportion of recurrent connections in the CA3 given by:

$$\text{syn}_i = \frac{7}{N_{\text{EC}}} = 0.014 \tag{1}$$

$$\text{syn}_0 = \frac{10}{N_{\text{CA3}}} = 0.01 \tag{2}$$

$$\text{syn}_{\text{CA3}} = \frac{50}{N_{\text{CA3}}} = 0.05 \tag{3}$$

where the numerator is the number of actual connections to the target layer, and the denominator is the total possible connections to the target layer equivalent to the number of neurons in the target layer.

Values of $syn_i = 0.014$, $syn_o = 0.01$, $syn_{CA3} = 0.1$ for all experiments unless otherwise stated.

2.3. Layer activations

We applied a k-winners-take-all (KWTA) rule in order to maintain sparse activation of neurons in each layer, thereby capturing the sparse activations and inhibitory connections observed in these regions [40]. We used PyTorch's topk to identify the k-largest activations. These top K neurons were set to 1 and the remainder of the neurons were set to 0. The proportion of neurons active in each layer, K, is given by:

$$K_{\text{EC}} = s_{\text{EC}} \cdot N_{\text{EC}} \tag{4}$$

$$K_{\text{DG}} = s_{\text{DG}} \cdot N_{\text{DG}} \tag{5}$$

$$K_{\text{CA3}} = s_{\text{CA3}} \cdot N_{\text{CA3}} \tag{6}$$

where s is the sparsity of activations (the proportion of active neurons in a layer). We used values of 0.1, 0.05 and 0.1 for s in K_{EC}, K_{DG} and K_{CA3}, respectively. We determine the layer activations, H, for the feedforward layers as follows:

$$H_l = \text{topk}(W_l H_{l-1} + b_l) \tag{7}$$

where *topk* is the KWTA function, W is the weights feeding into that layer, H_{l-1} are the previous layers activations, and b_l is the bias.

We calculate the recurrent output layer activations over two time steps as follows:

$$H_{\text{CA3}_k} = \text{topk}\left(H_{\text{CA3}_{k-1}} W_{\text{CA3}}\right) + H_{\text{DG}} W_o \tag{8}$$

where H_{CA3_k} is the CA3 activations at time k, where $H_{\mathrm{CA3}_{k-1}}$ is the CA3 activations at previous time, $WCA3$ are the recurrent weights, and $H_{DG}W_o$ is constant external input to the CA3 consisting of the DG activations and weights from the DG to CA3.

2.4. Weight updates

We randomly initialized the sparse synaptic weights representing neuronal connectivity using normal distributions with the variance proportional to input layer size, termed Xavier initialization [41]. Specifically, the variance, var, is given by the following:

$$\mathrm{var} = \frac{2}{n_{\mathrm{in}} + n_{\mathrm{out}}} \tag{9}$$

where n_{in} and n_{out} are the number of inputs and outputs of the layer, respectively.

The loss function, L, is given by the Hamming distance which measures the bit difference between two vectors, and is given by the following equation:

$$L = \sum_{i=1}^{k} |x_i - y_i| \tag{10}$$

where x and y are binary vectors representing the target and actual output patterns, and i represents the sequence of patterns.

To maintain syn_i, syn_o and syn_{ca3} to the preset levels, we generated a mask, m, such that any weights with zero values cannot change. Furthermore, weights are clamped to values between −1 and 1 by capping the values once they reach |1|. The weight update for our network is as follows:

$$W_i \leftarrow C\left(\left(W_i + \alpha\delta_{W_i}L\right)\right) \odot m_i \tag{11}$$

where W_i are the weights, α is the learning rate, $\delta_{Wi}L$ is the backpropagated error for the weights with respect to the loss, L, given above, m_i is the mask, and $C(\cdot)$ is the capping function given by:

$$C(x) = \begin{cases} 1 & x \geq 1 \\ x & -1 \leq x \leq 1 \\ -1 & x \leq -1 \end{cases} \tag{12}$$

The recurrent layer updates using backpropagation through time and the remainder of the layers learn through standard gradient descent (for review, see [42]) using the PyTorch Autograd module. Each update involves a complete pass of the full training set (not in batches).

Once the network generates an output pattern given an input pattern, we calculate the error between the target output and the generated output. Here, each type of input pattern is paired to one of four randomly generated target output patterns (i.e., A, B, C, or D output patterns)

(see 2.2 for the associative learning task). Using backpropagation, error signals generated at the output are sent back through the network layers to adjust the weights. This error signal is typically reduced in magnitude by a learning rate, a, typically a value less than one, which scales the relative size of weight updates at each iteration. We found a learning rate, a, of 0.0005 led to the least variable learning curves using a grid search (from 1×10^{-6} to 1, with 21 values spaced evenly on a log scale), testing on a separately generated validation set with 25 patterns. We used this value for all experiments. This procedure of determining the error, and propagating back through the network is repeated until 100 iterations have been reached for initial training, and 85 iterations for subsequent new or reversal training (see 2.2.1).

2.5. Measuring network performance

We tested the network's ability to reproduce the target output pattern upon presenting an input pattern from a test set of 25 patterns that the network had not seen during training. This accuracy measure is calculated using the Hamming distance, adjusted for the KWTA rule and sparsity level of the output layer:

$$\text{accuracy} = \left(1 - \frac{L}{K_{\text{CA3}}}\right) \times 100 \tag{13}$$

where L is the loss function given by the Hamming distance, as described in 2.1.3, and K_{CA3} is the number of neurons active at the given sparsity level for the output layer (K_{CA3} = 0.1 for all experiments). This accuracy measure is averaged across all test patterns to obtain an estimate of the network's performance.

2.6. Associative learning task

We trained the network on categorical input (EC) to output (CA3) pattern pairings. During each training stage, networks were presented an input pattern belonging to one of two categories (A and B, B and A, or C and D, see Section 2.2.1), and across training, networks learned to produce the associated target output pattern for that category. We generated input layer patterns drawn from four categories of input neurons, A, B, C, or D. The input layer is randomly divided into four subpopulations such that every input neuron belongs to either category A, B, C, or D. The total number of neurons active (i.e. neurons with activations set to 1) per input pattern is given by (K_{EC}) (see eq. 4).

Each input pattern from a single category consisted of active neurons of which 60% belong to the category, and the remaining 40% belong to the other three categories. For example, in A input patterns, of the 250 active neurons, 150 neurons belong to the A subpopulation, and 100 neurons belong to either B, C, or D subpopulations.

We randomly generated 50 of these input patterns for each category, comprising the training set. An additional 50 non-overlapping input patterns were generated (25 for testing, 25 for validation). We also generated 50 overlapping input patterns to control for the impact of overlap in the test set.

For each category, the input patterns are associated with a single category-specific output (Fig. 1). The overall goal of the training is for the network to identify which category the

majority of active EC neurons belong to and produce that category's target CA3 pattern. The CA3 target output patterns for the four categories are randomly generated producing four patterns with K_{CA3} active neurons.

2.7. Training stages

To create two distinct learning phases, $training_1$ and $training_2$, we only train on A and B patterns in $training_1$. Using backpropagation, we are able to use the error between the target output and the observed output patterns to train the weights of the network. Following $training_1$, we tested the network with or without first introducing post-training neurogenesis. The test consisted of presenting the network with new patterns from category A or B it had not seen, and measuring its ability to correctly produce the corresponding output pattern.

After $training_1$ on A and B categories, we then trained the network on a different set of patterns in $training_2$. We used either two new categories, C and D, (new learning) or the same categories with reversed outputs, B and A (reversal learning). In $training_2$, we measured both the accuracy attained once the performance plateaued (does not increase by more than 0.5% accuracy), and the number of epochs required to reach this final performance, t_{final}. These were performed with and without adding new neurons after $training_1$ (e.g. Fig. 2A).

2.8. Measuring forgetting and enhanced learning

We investigated the relationship and impact of neuronal addition and loss by measuring d_{forget} and d_{learn} given by:

$$d_{\text{forget}} = \frac{\left(\text{test}_{1_{\text{static}}} - \text{test}_{1_{\text{neurogenesis}}}\right)}{\text{test}_{1_{\text{static}}}} \tag{14}$$

$$d_{\text{learn}} = \frac{\left(\text{test}_{2_{\text{neurogenesis}}} - \text{test}_{2_{\text{static}}}\right)}{\text{test}_{2_{\text{static}}}} \tag{15}$$

$Test_1$ occurs after adding or not adding neurogenesis following $training_1$, and consists of testing the accuracy of the networks on the test sets for the A and B categories. $Test_2$ occurs after $training_2$ and involves testing the networks either with or without neurogenesis on the test sets used for $training_2$. Static refers to a network with no neurogenesis.

To avoid random effects of different initializations, for experiments correlating d_{forget} and d_{learn} across initializations, a copy of each model is made after $training_1$ before neurogenesis occurs. One copy receives neurogenesis and the other does not. These undergo $training_2$ to obtain d_{learn} and the comparison is made between neurogenic and static copies of the same model, repeated 500 times.

2.9. Neurogenesis

New neurons were exclusively added to the hidden layer (DG) of the model between $training_1$ and $training_2$ (see Section 2.2). These new neurons, similar to the 'mature' neurons, receive input from the EC and output onto the CA3 neurons, however, the sparsity of their connections were varied. New neurons also differ in terms of sparsity of activation (excitability) and learning rate (plasticity). We studied the impact of these properties on network performance by exploring the effects across the parameter space.

We calculated the KWTA for the 'mature' neurons as described in Section 2.1.2, and repeated this for the new neurons using the appropriate K values ($K_{\mathrm{DG_{mat}}} = 0.05$ for all experiments, and $K_{\mathrm{DG_{new}}} = 0.2$ for all unless otherwise stated). The resulting activations are summed to produce the first hidden state of the CA3 activations, Y_{CA3_1}:

$Y_{\mathrm{CA3}_1} = \left(Y_{\mathrm{DG}} W_{\mathrm{CA3_{mat}}}\right) + \left(Y_{\mathrm{DG}} W_{\mathrm{CA3_{new}}}\right)$ where $W_{\mathrm{CA3_{mat}}}$ and $W_{\mathrm{CA3_{new}}}$ are the weights connecting the 'mature' and new neurons, respectively, to the CA3. All weights for the new neurons were generated the same way weights were initialized in the 'mature' neurons (2.1.3).

2.10. Rate of neurogenesis

The level of neurogenesis, p_{new}, is the proportion of neurons added to the hidden layer. It is given by:

$p_{\mathrm{new}} = \frac{N_{\mathrm{new}}}{N_{\mathrm{DG}}}$ where N_{new} is the number of new neurons added in a single epoch, and N_{DG} is the number of neurons in the DG (hidden) layer.

For the majority of experiments we used $p_{new} = 0.05$. This was based on an analysis of effects of neurogenesis rates on differently sized neural networks. We found that $p_{new} = 0.05$ in a network with 5000 hidden units caused forgetting comparable to a much larger network (20000 hidden units) with $p_{new} = 0.002$, as the effects of new neurons scales with the size of hidden layer (see Figure 8).

Here, p_{new} corresponds to the number of new neurons that would be approximately 3–5 weeks of age at a given timepoint, which carry the unique properties of adult-generated granule cells that were explored in this paper.

2.11. Input and output connectivity

The impact of input and output connectivity of new neurons on network performance was explored. Similar to 'mature' neurons, a new neuron can receive a proportion, syn, of all possible connections, termed syn_i or syn_o for input and output synapses, respectively (see 2.1.1).

We varied connectivity of new DG neurons, syn_{new} for input and output connectivity using the following:

$$\text{syn}_{\text{new}_i} = \text{syn}_i \cdot c_i \text{ s.t. } c_i \sim \{0,\ 0.5,\ 1,\ 1.5,\ \ldots,\ 10\}$$

and

$$\text{syn}_{\text{new}_o} = \text{syn}_o \cdot c_o \text{ s.t. } c_o \sim \{0,\ 1,\ 2,\ \ldots,\ 10\}$$

where c_i and c_o are constant factors used to vary the input and output connnectivity, respectively. We measured accuracy after $training_1$ and $training_2$ with and without neurogenesis after $training_1$.

2.12. Plasticity

To explore the role of plasticity of newborn neurons on network performance, we modulated the learning rate, α_{new} (see 2.1.2) of only the new neurons, while keeping the learning rate of other neurons fixed across experiments. This is given by:

$$\alpha_{new} = \alpha_{mat} \cdot c_\alpha \text{ s.t. } c_\alpha \sim \{0,\ 0.5,\ 1,\ 1.5,\ \ldots,\ 20\}$$

where the c_α is the factor used to vary α_{new}.

A value of $c_\alpha = 4$ is used for all new/reversal learning experiments with higher values not appearing to provide any further benefit.

2.13. Excitability

Excitability was modelled in our networks as the sparsity of activations. The sparsity of the 'mature' DG, $n_{\text{DG}_{\text{mat}}}$ was 0.05 for all experiments, i.e. with 250 of the 5000 DG neurons active on any given forward pass, based on the KWTA-rule (see 2.1.2). To explore the role of excitability of newborn neurons on network performance, we modulated the excitability of only the new neurons, while keeping the excitability (i.e. sparsity) of the 'mature' neurons fixed. This is given by:

$$n_{\text{DG}_{\text{new}}} = n_{\text{DG}_{\text{mat}}} \cdot c_{\text{exc}} \text{ s.t. } c_{exc} \sim \{0,\ 0.5,\ 1,\ 1.5,\ \ldots,\ 20\}$$

A value of $c_{exc} = 4$ such that $n_{\text{DG}_{\text{new}}} = 0.2$ was used for all other experiments, i.e. 50 of 250 new neurons were active in each forward pass. The sparsity of new neurons was chosen to be 4 times greater based on *in vivo* results.

2.14. Neuronal and synaptic turnover

In some experiments, we not only added neurons, but explored the loss of neurons as well. The ratio of loss to addition, $p_{replace}$ is defined by

$$p_{replace} = N_{loss} / N_{new}$$

where N_{loss} is the number of 'mature' neurons lost, and N_{new} is the number of new neurons added.

Neuronal turnover is defined here as a loss of neurons when new neurons are added. We tested the impact of turnover by varying N_{loss} and N_{new} systematically, from values of 0 to 1000, at increments of 25. Due to the computational intensity of this approach, we decided to stop at a maximum of 20% loss or addition to the DG population, but predict that this becomes biologically irrelevant at values beyond this, given that rates of neurogenesis observed *in vivo* are much lower.

2.15. Statistical analyses and plotting

All statistics were performed using the scipy.stats module in Python 3.6. Error bars on graphs represent the standard error of the mean, across different initializations of the model, where each experiment is repeated 20 times unless otherwise stated. Comparisions were made using a two-way factorial, or one-way ANOVA with Tukey's HSD post hoc test where appropriate. Significance indicated by an asterisk (*) for p-values less than 0.05 unless otherwise stated. Graphs were generated using the *seaborn* package in Python 3.6, and compiled into figures using Inkscape.

3. Results

3.1. Generation of a 3 layer ANN with neurogenesis occurring in the hidden layer

To study the impact of neuron addition on memory stability we generated a 3 layer ANN, with the input layer corresponding to the EC, the hidden layer corresponding to the DG, and the output layer corresponding to CA3, which has dense recurrent connections (Fig. 1A). The network is comprised of 500, 5000 and 1000 units in the EC, DG and CA3 respectively, capturing relative neuron number ratios in these regions in rodents [36]. Neurogenesis occurred only in the hidden (DG) layer. Mature DG neurons receive sparse inputs from the EC [45], and send sparse output connections to CA3 [36], and are sparsely activated relative to EC and CA3 [46]. In contrast, new DG neurons receive fewer inputs from the EC [16], send more output connections to the CA3 [47], and are more excitable relative to their mature counterparts [12]. We incorporated these distinct wiring and physiological properties of new and mature DG neurons in our model.

We trained this network on a categorization task. Our training set comprised of 50 A and 50 B binary input patterns (with sparsity of 10%) (Fig. 1B). These patterns overlapped by 40% in cluster membership. All A or B patterns were associated with a single A or B output pattern, respectively, each with a sparsity of 10% of the output layer size. A further 40 patterns were generated for A and B, with no overlapping neurons, comprising our test set (see 2.2 for further details). We confirmed that our model is able to categorize A and B input patterns significantly above chance levels of 10% (Fig. 1C). We further established that categorization performance did not depend upon selecting only non-overlapping neurons in the test set: Testing on patterns with the same overlap as the training set (i.e., 40%) produced similar categorization accuracy (Supplemental Fig 1).

A typical experiment involved training on A and B patterns, and then adding new neurons (neurogenesis condition) or not (static condition) to the middle layer of the network. The network was then tested on the remaining A and B patterns (Fig. 1D). A training epoch consisted of presenting our input patterns to the model, calculating the Hamming distance between the expected output A or B patterns and the actual output patterns from our model (Fig. 1E), and propagating the error backwards through the network using stochastic gradient descent (see 2.1.3).

3.2. Hippocampal neurogenesis induces forgetting of stored memories

We first assessed how neurogenesis impacts the stability of stored memories. To do this we trained the network in the AB categorization task, and then added new neurons to the hidden layer (Fig. 2A). We varied the proportion of new neurons added, where proportion is the ratio between new and old neurons ($p_{new} = N_{new}/N_{DG}$). After adding new neurons, we compared performance on previously unseen AB test patterns. We hypothesized that post-learning neurogenesis would cause forgetting, with more forgetting as more neurons are added to the DG. Consistent with our hypothesis, we found that forgetting was dependent on the proportion of new neurons added, with forgetting plateauing (when the change in performance is less than 5% between epochs) when $p_{new} > 0.05$ (Fig. 2B). This corresponds approximately to levels of neurogenesis observed in young adult rodents across a 2–3 week time-span, and based on cross-species analyses of rates of neurogenesis, this value likely also falls within the rates observed in primates and humans [1]. Therefore, we chose this rate of neurogenesis in all future experiments.

To further investigate how forgetting may be further modulated by initial encoding, we attempted to curb forgetting by overtraining our network by training for more epochs, or providing our network with more input pattern exemplars during training (Fig. 2C,D). In both cases, we saw similar forgetting effects despite making the memories stronger.

3.3. Rates of forgetting are modulated by connectivity of newborn neurons

As newborn neurons integrate into hippocampal circuits they form input and output connections. We hypothesized that forgetting should be sensitive to the extent of integration of newborn neurons, such that increased input or output connectivity should produce more forgetting. Therefore, we next examined the impact of different levels of new neuron input and output connectivity on forgetting. To do this we trained our network on the AB categorization task as before, and, following training, new neurons were added at the p_{new} rate (i.e., 0.05) (Fig. 3A). We first manipulated the input connectivity of new neurons (relative to mature granule cell connectivity), while maintaining output connections at mature granule cell levels. We initially selected 3 levels of input connectivity for newborn neurons: hypo-integration (corresponding to 0.5X mature granule cell input connectivity), normal-integration (matching mature granule cell input connectivity) and hyper-integration (corresponding to 4X mature granule cell input connectivity). We found that all levels of input integration resulted in similar levels of forgetting (Fig. 3B left). We then comprehensively evaluated forgetting across a range of input connectivities. We found that there was no effect of increasing input connectivity of new neurons on forgetting within the parameter range of 1 to 10 times the connectivity of a mature neuron (Fig. 3C).

We next manipulated the output connectivity of new neurons (relative to mature granule cell connectivity), while maintaining input connections at mature granule cell levels. We selected 3 levels of output connectivity for newborn neurons: hypo-integration (corresponding to 0.5X mature granule cell output connectivity), normal-integration (matching mature granule cell output connectivity) and hyper-integration (corresponding to 4X mature granule cell output connectivity). With normal output connectivity, neurogenesis induced forgetting, as we observed previously. These levels of forgetting were exacerbated by increasing output connections of newborn neurons, while restricting the integration of newborn neurons into the circuit abolished the forgetting effect (Fig. 3B right). We then comprehensively evaluated forgetting across a range of output connectivities. We found that there was an effect of increasing output connectivity of new neurons on forgetting within the parameter range of 1 to 10 times the connectivity of a mature neuron (Fig. 3D).

Finally, we changed input and output connectivity of newborn neurons simultaneously. We observed a similar increase in forgetting with increased connectivity, as seen with manipulating output connectivity alone. This lack of interaction suggests that forgetting is primarily sensitive to changes in output connectivity of new neurons (Fig. 3E).

3.4. Rates of forgetting are modulated by excitability of newborn neurons

Newborn neurons are transiently more excitable than mature granule cells between 4–7 weeks of age [12,13]. We therefore next examined how new neuron excitability modulates forgetting in our ANN. To do this we trained our network on the AB categorization task as before, and, following training, new neurons were added with varying excitability levels (Fig. 4A). In our model excitability was determined the sparsity of activation (i.e., using the KWTA-rule, see 2.1.2). To mimic excitability changes, we varied sparsity in the new neuron population relative to sparsity in the mature neuron population, with increased excitability corresponding to decreased sparsity (Fig. 4B). We hypothesized that increased excitability of new neurons (relative to their mature counterparts) would cause more forgetting. As sparsity decreased (i.e., excitability increased) in the new neuron population, forgetting increased (Fig. 4C,D). Forgetting was also observed even when sparsity in the new neuron population was higher (i.e., excitability decreased relative to the mature granule cell population) (Fig. 4C). More importantly, when fold-changes in excitability matched those typically observed *in vivo* (i.e., ~2–4X [15,43]), forgetting rates were near maximal. We chose a value of 4X the excitability of the mature neurons as the default value for all following experiments.

3.5. Forgetting facilitates reversal learning

Forgetting stored memories may be useful in environments that are dynamic, and where contingencies are constantly changing [48]. Consistent with this idea, we have shown that artificially elevating hippocampal neurogenesis weakens established memories, but facilitates encoding of new information that conflicts with previous learning [24]. Therefore, we next explored how forgetting impacts new category learning that is either conflictual or non-conflictual in nature. We hypothesized that neurogenesis-mediated forgetting should positively correlate with levels of conflictual, but not non-conflictual, learning.

To do this we trained the network on the AB categorization, and then added new neurons to the hidden layer and tested performance on the AB categorization (test 1). Next we re-trained the network on a reversal task, where inputs pattern which would have normally corresponded to A now corresponding to B (and vice-versa). We tested performance on the BA categorization to assess the efficiency of reversal learning in networks with and without neurogenesis (test 2) (Fig. 5A).

Post-training neurogenesis induced forgetting of the AB categorization in test 1 (inset, Fig. 5B). Across 500 replications we calculated d_{forget} (i.e., the difference between performance in static vs neurogenesis conditions). As expected, networks with neurogenesis outperformed static networks in the reversal task (i.e., BA categorization). BA learning was faster in networks with neurogenesis, reaching asymptotic performance far sooner than static networks (Fig 5B). To quantify this advantage we computed d_{learn} (i.e., the difference in performance between neurogenic and static networks when the neurogenic network performance reaches asymptote). To assess how forgetting interacts with subsequent conflictual learning we plotted d_{learn} against d_{forget} (Fig. 5C). We found that forgetting moderately predicted subsequent conflictual learning, consistent with published mouse behavioral experiments [24].

We additionally assessed non-conflictual learning (i.e., new category learning (CD)) following forgetting (Fig. 5D). As before, post-training neurogenesis induced forgetting of the AB categorization in test 1 (Fig 5D, inset). However, there was only modest facilitation of CD learning in the neurogenic network compared to the static network (Fig. 5D), suggesting that forgetting primarily facilitates subsequent learning in conflictual (but not non-conflictual) situations. Consistent with this, there was no relationship between d_{learn} and d_{forget} (Fig 5E).

3.6. Excitability and output hyperintegration drive reversal learning enhancements

To test the impact of plasticity, excitability, and input and output connectivity on the reversal learning enhancement, we varied these parameters in new neurons relative to the values used for mature neurons. We trained the network on the AB categorization, added new neurons with properties varied across the parameter space, and retrained in the reversal categorization (BA), as before (Fig. 6A). We hypothesized that increasing plasticity (i.e., learning rate), excitability, and input and output connectivity would enhance reversal learning (Fig 6B–E). We found that reversal learning accuracy was sensitive to increases in excitability (Fig. 6C) and output connectivity (Fig. 6E). However, reversal learning accuracy was largely insensitive to increases in learning rate (or plasticity) (Fig. 6B) and changes in input connectivity (Fig. 6D,F).

3.7. Forgetting and reversal learning in networks with loss and addition of new neurons (turnover)

Although the dentate gyrus is thought to have net growth through early to mid-adulthood in mice [49], this doesn't exclude the possibility that fully mature granule cells are lost. Indeed, in rats ~17% of postnatally generated granule cells die between 2–6 months [50], indicating that the addition of new neurons is accompanied by loss of some fully mature

granule cells. We next explored how forgetting and reversal learning are affected by both the loss of mature granule cells, and the addition of new neurons in our ANN.

We first examined forgetting. We trained our network on the AB categorization, and then varied rates of neuron loss and addition and computed d_{forget}. Compared to the static network, neuronal addition (in the absence of neuronal loss) increased forgetting, as seen previously (Fig. 7A). This increase occurred with relatively low rates of neuron addition (i.e., ~2%). Similarly, compared to the static network, neuronal loss (in the absence of neuron addition) increased forgetting. However, this only occurred when rates of neuronal loss exceeded ~10%, and forgetting rates were much lower compared to those observed with only neuronal addition. This suggests that forgetting in this model is more sensitive to neuronal addition than neuronal loss.

We next explored why neuron addition produced greater levels of forgetting than neuronal loss. In our model, neuronal loss is random and so it is unlikely that all neurons with the highest information content (i.e., neurons that pass the KWTA threshold) would be lost. Furthermore, when high information content neurons are lost, they would be most likely replaced by the other neurons with high information content (i.e., neurons that were just below the KWTA threshold) and therefore little forgetting would be predicted. We tested this intuition in the following way. We trained the network as before and then silenced the top k neurons. Then we either allowed replacement from the next ranking cohort of neurons or the bottom ranking cohort of neurons. We found that replacement with the next ranking cohort led to little forgetting. In contrast, there was much higher levels of forgetting when replacement neurons were selected from the bottom ranking cohort (Supplemental Figure 2). This suggests the network is resilient to neuron loss because replacement neurons are drawn from the next ranking neurons which will inevitably contain rich information.

In the same model, we next evaluated reversal learning following forgetting (i.e., d_{learn}). Compared to the static network, neuronal addition (in the absence of neuronal loss) increased d_{learn}, and this facilitation was observed at relatively low rates of neurogenesis (> ~2%) (Fig. 7B). In contrast, neuronal loss (in the absence of neuronal addition) impaired reversal learning only when loss exceeded ~10% of total DG. However, neuronal addition appears to be protective against the effects of neuronal loss on reversal learning. At rates of neuronal loss > ~10%, reversal learning was still enhanced when neurogenesis levels were above ~2%.

3.8. Forgetting and facilitated reversal learning at low rates of neurogenesis is observed in large networks

In our network, the input (EC), middle (DG) and output (CA3) layer had 500, 5000, 1000 neurons, respectively. To evaluate whether network size impacts observed forgetting rates following neurogenesis, we trained and tested our network as before (Fig. 8A) and varied numbers of neurons in input, middle and output layers while maintaining the 1:10:2 ratio. Forgetting was observed for all networks containing > 500 DG units, indicating that this is a robust phenomenon, even in very large networks (i.e., containing 20,000 DG units) (Fig. 8B). Moreover, the minimum proportion of neurons added in order to observe forgetting (defined as a minimum of 10% decrease in network performance) decays as network size

increases (Fig. 8C), suggesting that even lower rates of neurogenesis may significantly impact memory storage at more biologically relevant network sizes (e.g. 1,200,000 neurons in rat DG and 7,000,000 in macaques [35]).

We next explored the difference between models where new neurons have the same properties as their mature counterparts, and models where new neurons have the distinct properties observed biologically (increased excitability, different wiring etc.). We found that by distinguishing the new neurons from the mature neurons by changing their properties to simulate those observed biologically, we see that even lower levels of neurogenesis are required to observe similar effects (0.2%) which is more in line with estimates of human rates of neurogenesis [38] (Fig. 8B,C).

To evaluate whether enhanced reversal learning is observed across both smaller and larger (and more biologically plausible) ANNs, we varied the size of the DG (and scaled input and output layers accordingly). We trained the network on the AB categorization, added new neurons (or not), and then retrained in the reversal categorization (BA), as before. We found that reversal learning depended upon network size. Whereas in small networks (e.g., < 5,000 neurons in the hidden layer) forgetting occurred with low to moderate rates of neurogenesis, this was not associated with enhanced reversal learning. Only when networks exceeded ~5,000 neurons in the hidden layer, was enhanced reversal learning observed. Indeed, reversal learning increased as a function of network size (at least up to hidden layers containing 20,000 neurons) (Figure 8D, E).

We again explored the difference between models where new neurons have the same or different properties than their mature counterparts. We found that by distinguishing the new neurons from the mature neurons, we also observed that lower levels of neurogenesis were required to observe similar learning enhancements (defined as a minimum of a 10% enhancement in network performance) (Fig. 8D right, and 8E blue line).

4. Discussion

4.1. Overview of results

In this paper, we built a model of the hippocampus to explore the impact of ongoing neurogenesis on memory using a categorization task. Consistent with previous models [30,51], we found that post-training neurogenesis caused forgetting. These forgetting effects were observed regardless of network size (from 500–20,000 DG neurons), and when the network was over-trained (either by providing the network with more exemplars during training or training over more epochs). We showed that forgetting was modulated by properties of new neurons, and identified output connectivity and excitability as key features of new neurons that promote forgetting. We also found that forgetting facilitated conflictual learning. When learning patterns that conflicted with previous training sets, higher rates of forgetting of the original categories was associated with more efficient reversal learning. Enhanced reversal learning was similarly modulated by the output connectivity and excitability of new neurons, and was observed regardless of network size. These findings support empirical work that shows that hippocampal neurogenesis regulates forgetting [22,24,52–54], and that partial forgetting allows animals to behave more flexibly in dynamic

environments [24], where contingencies change with time (see also [48]). Importantly the effects of neurogenesis on memory stability and flexibility were observed at biologically plausible rates of new neuron addition (e.g., when new neuron addition corresponded to ~0.2% of DG layer size).

4.2. Exploration of parameters

Exploring properties of new neurons in isolation allowed us to examine their distinct contributions to forgetting and subsequent learning (either new or reversal learning). We found that input connectivity had little impact on either forgetting or reversal learning. In contrast, changing output connectivity markedly impacted memory stability and associated reversal learning. Similarly, increasing excitability also promoted forgetting and enhanced reversal learning. While excitability has been related to temporal integration (across multiple episodes) [33], our findings here point to an additional role in regulation of memory stability. Finally, the range of learning rates that we used in our model were lower than typically explored in conventional machine learning applications. We did not observe any relationship between learning rate of new DG neurons and reversal learning performance, most likely because we examined this more restricted learning rate range.

Our neurogenic networks were better than static networks in the reversal learning task (where conflictual patterns are learned). This corresponds with previous work that found post-training increases in hippocampal neurogenesis facilitated conflictual learning in mice [24]. We also found that levels of forgetting were weakly, but significantly, correlated with subsequent learning in the reversal (or conflictual) learning, and this same relationship was observed in mice [24]. Since reversal learning is useful in non-static environments, this supports the notion that hippocampal neurogenesis is important for cognitive flexibility [48,55]. We also observed improved learning of new patterns (i.e., CD categorization), consistent with previous models that emphasize the impact of neurogenesis on encoding new information [29,31–34,56]. However, compared to the effects on forgetting and reversal learning, this effect was modest in magnitude. This is most likely because large networks (such as the one we used) can learn the relatively small number of simple categorizations quite efficiently. Future work testing the ability of neurogenesis to enhance performance on more difficult learning tasks may address this issue.

Forgetting in these types of models might be caused by new neuron addition, loss of mature neurons or some combination of both. In the adult rodent hippocampus, it is likely that both new neuron addition, and loss of mature granule cells occur concurrently. Indeed, in line with this idea, there is significant loss of developmentally-generated neurons (from P6) between 2 and 6 months of age in rats [50]. To investigate this, we incorporated both neuron addition and neuron loss in our model. Surprisingly, we found that forgetting was primarily driven by new neuron addition: In order to produce equivalent forgetting, neuron loss needed to be an order of magnitude higher than neuron addition.

These results suggest that the hippocampus is quite resilient to neuron loss. We suggest that three features of the network likely contribute to its sensitivity to neuron addition vs. neuron loss. First, network resiliency to neuron loss is most likely because replacement neurons are drawn from the next ranking neurons which also contain rich information. When we

forced the network to use low ranking cohorts of neurons for replacement then significant forgetting was observed (Supplemental Figure 2). Second, the output layer is a recurrent network. This allows our network to pattern complete, and therefore even impoverished DG inputs may be sufficient for successful recall. Third, for neuron addition, the new neurons do not contain any information (because they are randomly initialized) and therefore these new cohorts add only noise, which may overwhelm the ability of the recurrent network to pattern complete.

4.3. Comparison to other neurogenesis models

Other models have also examined the impact of neurogenesis on forgetting of hippocampus-dependent memories [30,51]. These models have also found that neurogenesis promotes forgetting, but, for the most part, used much higher rates of neurogenesis (e.g., 5–30%). These rates exceed values reported in the human hippocampus [1–4] and the hippocampus of non-human mammalian species [6–10]. In contrast, we observed forgetting when the new neurons represented as little as 0.2% of the total DG layer. We suspect that network size may have contributed to the prevalence of forgetting at proportionally much lower rates of neurogenesis in our study. In our main network, the middle layer contained 5000 DG neurons, whereas previous models contained only 10–1000 [28,30]. When we systematically explored the impact of network size on forgetting rates, in small networks (500–1000 DG neurons), much higher rates of neurogenesis were similarly required to produce forgetting (~5–10%). Consistent with this, in other models with fewer DG neurons, much higher rates of neurogenesis rates were necessary to produce forgetting (5–30%) [30,51]. In contrast, the levels of neurogenesis where we observe forgetting in our model match those observed in adult rodents (e.g., 1–5% per month [37,57]), and estimated in humans (0.12% per month [38]).

Our model additionally differed in a number of other important ways from previous models. First, we maintained the approximate ratio of neurons within the EC → DG → CA3 trisynaptic circuit. Second, in our model the DG was sparsely connected, and we imposed sparse activity in this layer, matching sparse activation of the dentate that is observed *in vivo* [43,44]. The sparse connectivity additionally allowed us to explore how varying levels of integration of newborn neurons impacted forgetting and conflictual learning; something that was not possible in previous models that were fully-connected (see below). Third, we incorporated recurrency in the output layer. Recurrency allows for pattern completion, even with impoverished inputs, and mimics the autoassociative architecture of CA3 [58,59]. Despite recurrency making the network more resistant to changing inputs (i.e., resistant to forgetting), we still observed significant forgetting at low rates of neurogenesis, and even after overtraining.

Previous studies have established that new neurons transiently differ from their developmentally generated neighbors. For example, they receive fewer input connections from the EC, and send more output connections to the CA3 than their mature counterparts. Moreover, they are highly excitable (they have a greater probability of firing) [11,12] and also more plastic (are more readily able to undergo LTP) [60,61]. One of the core advantages of our model was that we were able to systematically explore how each of these

features contribute to forgetting. We found that neurogenesis alone (without incorporating the properties of new neurons) causes forgetting and enhanced reversal learning. However, building the properties of new neurons into our model further exacerbated these effects.

4.4. Limitations of the model

4.4.1. Backpropagation learning algorithm and multi-trial learning task—

Theoretical accounts of the role of the hippocampus emphasize its role in episodic learning, in which experiences are rapidly encoded in 'one shot' [62], without the need for multiple trials. In contrast, the learning algorithm used in our model required multiple exposures to the input-output pairings, and therefore learning was much more gradual in nature and required many trials. We note, however, that some forms of hippocampus-dependent learning are more incremental in nature, and perhaps related to the type of learning algorithm we use here. This includes many forms of maze learning (e.g., water maze [63], radial arm maze [64], Barnes maze [65,66]), non-spatial mapping (e.g., continuous tonal maps [67]), category learning (e.g., visual categorization [68]) and paired associate learning [69]. Indeed, post-training elevation of hippocampal neurogenesis induces forgetting of single-trial learning memories (e.g., contextual fear conditioning) [22,53], as well as multi-trial learning memories (e.g., water maze [22,24], PAL [Epp et al, this issue]).

4.4.2. Implementation of neurogenesis—

In our model, neurogenesis was implemented as a one-time addition of new neurons after learning AB categories, but before learning new (CD) or conflicting (BA) categories. However, it is likely that there is an impact of neurogenesis across time, such that a heterogeneous population of immature neurons born at different timepoints, and thus different stages of maturity, may support memory in more nuanced ways than captured in our model. Related to this issue, previous models concluded that neurogenesis protects, rather than overwrites, old memories [31–33]. In these models, neurogenesis was steady-state (i.e., occurred continuously) as the network encoded successive memories. Moreover, the properties of new neurons depended on their age−/− younger new neurons were transiently more plastic, but plasticity levels declined as neurons aged. Accordingly, memories that occurred close together in time were encoded by overlapping populations of highly plastic, new neurons, but memories that occurred further apart in time were encoded in separate populations of highly-plastic, new neurons. This allowed new learning to occur without impacting old learning. In contrast, in our network, we modeled a situation where there was a net gain in neurogenesis levels following learning. As discussed above, we suggest that new neurons added noise to the network and the degraded already stored information. Consistent with our findings, when one of these previous models implemented neurogenesis as a net gain (rather than steady-state), forgetting of previously learned information was also observed (see Fig. 5E–G, [33]).

5. Conclusions

Low rates of hippocampal neurogenesis have been reported in humans and non-human mammalian species [1]. Despite these low rates hippocampal neurogenesis, manipulations of hippocampal neurogenesis in rodents have consistently altered cognitive function, including forgetting. Our findings here showing that forgetting and conflictual learning are impacted

by relatively low numbers of new neurons (especially in larger networks), is consistent with this perspective. Moreover, these low rates of hippocampal neurogenesis have a greater effect when incorporating the unique physiological and anatomical properties of immature adult-generated neurons, such as enhanced plasticity, excitability and hyperintegration. We predict that at the scale of biological networks (~1,200,000 neurons in the rodent DG), small changes in hippocampal neurogenesis are sufficient to profoundly impact cognitive function.

We would like to acknowledge the following open source software used for building the model, and subsequent data analysis and visualization:

- PyTorch: Paszke A. et al. Automatic differentiation in PyTorch (2017), NIPS-W. (http://pytorch.org/)
- SciPy: Jones E., Oliphant E., Peterson P.,et al.SciPy: Open Source Scientific Tools for Python(2001). (https://www.scipy.org/)
- Numpy: OliphantE.A.,Guide to NumPy, USA: Trelgol Publishing, (2006). (http://www.numpy.org/)
- Matplotlib: Hunter J. D.,Matplotlib: A 2D graphics environment, Computing In Science & Engineering(2007). (https://matplotli-b.org/)
- Scikit-learn:Pedregosa,et al.,Scikit-learn: Machine learning in Python, JMLR12, 2825–2830(2011). (http://scikit-learn.org/stable/)
- Pandas: McKinney W.Data Structures for Statistical Computing in Python, Proceedings of the 9th Python in Science Conference, 51–56 (2010). (https://pandas.pydata.org/)
- IPython: Perez F., Granger B.E.IPython: A System for Interactive Scientific Computing,Computing in Science & Engineering,9, 21–29 (2007). (https://ipython.org/)
- Seaborn: https://seaborn.pydata.org/
- Inkscape, Inkscape Project (2011): https://inkscape.org

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work was supported by CIHR Foundation grant to PWF (FDN143227), and a Canadian Institute for Advanced Research (CIFAR) grant to SAJ, BAR and PWF. LT was supported by fellowships from NSERC and The Hospital for Sick Children (Restracomp program).

References

[1]. Snyder JS, Recalibrating the relevance of adult neurogenesis, Trends Neurosci. 42 (2019) 164–178. [PubMed: 30686490]

[2]. Boldrini M, Fulmore CA, Tartt AN, Simeon LR, Pavlova I, Poposka V, Rosoklija GB, Stankov A, Arango V, Dwork AJ, Hen R, John Mann J, Human Hippocampal Neurogenesis Persists

throughout Aging, Cell Stem Cell. 22 (2018) 589–599, 10.1016/j.stem.2018.03.015 e5. [PubMed: 29625071]

[3]. Eriksson PS, Perfilieva E, Björk-Eriksson T, Alborn AM, Nordborg C, Peterson DA, Gage FH, Neurogenesis in the adult human hippocampus, Nat. Med 4 (1998) 1313–1317. [PubMed: 9809557]

[4]. Moreno-Jiménez EP, Flor-García M, Terreros-Roncal J, Rábano A, Cafini F, Pallas-Bazarra N, Ávila J, Llorens-Martín M, Adult hippocampal neurogenesis is abundant in neurologically healthy subjects and drops sharply in patients with Alzheimer's disease, Nature Med. (2019), 10.1038/s41591-019-0375-9.

[5]. Sorrells SF, Paredes MF, Cebrian-Silla A, Sandoval K, Qi D, Kelley KW, James D, Mayer S, Chang J, Auguste KI, Chang EF, Gutierrez AJ, Kriegstein AR, Mathern GW, Oldham MC, Huang EJ, Garcia-Verdugo JM, Yang Z, Alvarez-Buylla A, Human hippocampal neurogenesis drops sharply in children to undetectable levels in adults, Nature 555 (2018) 377–381. [PubMed: 29513649]

[6]. Lowe A, Dalton M, Sidhu K, Sachdev P, Reynolds B, Valenzuela M, Neurogenesis and precursor cell differences in the dorsal and ventral adult canine hippocampus, Neurosci. Lett 593 (2015) 107–113. [PubMed: 25778416]

[7]. Jabès A, Lavenex PB, Amaral DG, Lavenex P, Postnatal development of the hippocampal formation: a stereological study in macaque monkeys, J. Comp. Neurol 519 (2011) 1051–1070. [PubMed: 21344402]

[8]. Ngwenya LB, Heyworth NC, Shwe Y, Moore TL, Rosene DL, Age-related changes in dentate gyrus cell numbers, neurogenesis, and associations with cognitive impairments in the rhesus monkey, Front. Syst. Neurosci 9 (2015) 102. [PubMed: 26236203]

[9]. Altman J, Das GD, Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats, J. Comp. Neurol 124 (1965) 319–335. [PubMed: 5861717]

[10]. Kempermann G, Early determination and long-term persistence of adult-generated new neurons in the hippocampus of mice, Development 130 (2003) 391–399, 10.1242/dev.00203. [PubMed: 12466205]

[11]. Danielson NB, Kaifosh P, Zaremba JD, Lovett-Barron M, Tsai J, Denny CA, Balough EM, Goldberg AR, Drew LJ, Hen R, Losonczy A, Kheirbek MA, Distinct Contribution of Adult-Born Hippocampal Granule Cells to Context Encoding, Neuron 90 (2016) 101–112. [PubMed: 26971949]

[12]. Marín-Burgin A, Mongiat LA, Pardi MB, Schinder AF, Unique processing during a period of high excitation/inhibition balance in adult-born neurons, Science 335 (2012) 1238–1242. [PubMed: 22282476]

[13]. Dieni CV, Panichi R, Aimone JB, Kuo CT, Wadiche JI, Overstreet-Wadiche L, Low excitatory innervation balances high intrinsic excitability of immature dentate neurons, Nat. Commun 7 (2016) 11313. [PubMed: 27095423]

[14]. Ge S, Yang C-H, Hsu K-S, Ming G-L, Song H, A critical period for enhanced synaptic plasticity in newly generated neurons of the adult brain, Neuron 54 (2007) 559–566. [PubMed: 17521569]

[15]. Schmidt-Hieber C, Jonas P, Bischofberger J, Enhanced synaptic plasticity in newly generated granule cells of the adult hippocampus, Nature 429 (2004) 184–187. [PubMed: 15107864]

[16]. Toni N, Matthew Teng E, Bushong EA, Aimone JB, Zhao C, Consiglio A, van Praag H, Martone ME, Ellisman MH, Gage FH, Synapse formation on neurons born in the adult hippocampus, Nature Neurosci. 10 (2007) 727–734, 10.1038/nn1908. [PubMed: 17486101]

[17]. Restivo L, Niibori Y, Mercaldo V, Josselyn SA, Frankland PW, Development of adult-generated cell connectivity with excitatory and inhibitory cell populations in the hippocampus, J. Neurosci 35 (2015) 10600–10612. [PubMed: 26203153]

[18]. Drew LJ, Kheirbek MA, Luna VM, Denny CA, Cloidt MA, Wu MV, Jain S, Scharfman HE, Hen R, Activation of local inhibitory circuits in the dentate gyrus by adult-born neurons, Hippocampus 26 (2016) 763–778. [PubMed: 26662922]

[19]. McAvoy K, Besnard A, Sahay A, Adult hippocampal neurogenesis and pattern separation in DG: a role for feedback inhibition in modulating sparseness to govern population-based coding, Front. Syst. Neurosci 9 (2015) 120. [PubMed: 26347621]

[20]. Clelland CD, Choi M, Romberg C, Clemenson GD Jr., Fragniere A, Tyers P, Jessberger S, Saksida LM, Barker RA, Gage FH, Bussey TJ, A functional role for adult hippocampal neurogenesis in spatial pattern separation, Science 325 (2009) 210–213. [PubMed: 19590004]

[21]. Nakashiba T, Cushman JD, Pelkey KA, Renaudineau S, Buhl DL, McHugh TJ, Rodriguez Barrera V, Chittajallu R, Iwamoto KS, McBain CJ, Fanselow MS, Tonegawa S, Young dentate granule cells mediate pattern separation, whereas old granule cells facilitate pattern completion, Cell 149 (2012) 188–201. [PubMed: 22365813]

[22]. Akers KG, Martinez-Canabal A, Restivo L, Yiu AP, De Cristofaro A, Hsiang H-LL, Wheeler AL, Guskjolen A, Niibori Y, Shoji H, Ohira K, Richards BA, Miyakawa T, Josselyn SA, Frankland PW, Hippocampal neurogenesis regulates forgetting during adulthood and infancy, Science 344 (2014) 598–602. [PubMed: 24812394]

[23]. Burghardt NS, Park EH, Hen R, Fenton AA, Adult-born hippocampal neurons promote cognitive flexibility in mice, Hippocampus 22 (2012) 1795–1808, 10.1002/hipo.22013. [PubMed: 22431384]

[24]. Epp JR, Mera RS, Köhler S, Josselyn SA, Frankland PW, Neurogenesis-mediated forgetting minimizes proactive interference, Nature Commun. 7 (2016), 10.1038/ncomms10838.

[25]. Kitamura T, Inokuchi K, Role of adult neurogenesis in hippocampal-cortical memory consolidation, Mol. Brain 7 (2014) 13. [PubMed: 24552281]

[26]. Snyder JS, Soumier A, Brewer M, Pickel J, Cameron HA, Adult hippocampal neurogenesis buffers stress responses and depressive behaviour, Nature 476 (2011) 458–461. [PubMed: 21814201]

[27]. Santarelli L, Saxe M, Gross C, Surget A, Battaglia F, Dulawa S, Weisstaub N, Lee J, Duman R, Arancio O, Belzung C, Hen R, Requirement of hippocampal neurogenesis for the behavioral effects of antidepressants, Science 301 (2003) 805–809. [PubMed: 12907793]

[28]. Meltzer LA, Yabaluri R, Deisseroth K, A role for circuit homeostasis in adult neurogenesis, Trends Neurosci. 28 (2005) 653–660. [PubMed: 16271403]

[29]. Chambers RA, Andrew Chambers R, Potenza MN, Hoffman RE, Miranker W, Simulated Apoptosis/Neurogenesis Regulates Learning and Memory Capabilities of Adaptive Neural Networks, Neuropsychopharmacology 29 (2004) 747–758, 10.1038/sj.npp.1300358. [PubMed: 14702022]

[30]. Weisz VI, Argibay PF, Neurogenesis interferes with the retrieval of remote memories: forgetting in neurocomputational terms, Cognition 125 (2012) 13–25. [PubMed: 22841299]

[31]. Wiskott L, Rasch MJ, Kempermann G, A functional hypothesis for adult hippocampal neurogenesis: avoidance of catastrophic interference in the dentate gyrus, Hippocampus 16 (2006) 329–343. [PubMed: 16435309]

[32]. Finnegan R, Becker S, Neurogenesis paradoxically decreases both pattern separation and memory interference, Front. Syst. Neurosci 9 (2015) 136. [PubMed: 26500511]

[33]. Aimone JB, Wiles J, Gage FH, Computational influence of adult neurogenesis on memory encoding, Neuron 61 (2009) 187–202. [PubMed: 19186162]

[34]. DeCostanzo AJ, Fung CCA, Fukai T, Hippocampal neurogenesis reduces the dimensionality of sparsely coded representations to enhance memory encoding, Front. Comput. Neurosci 12 (2018) 99. [PubMed: 30666194]

[35]. Amaral DG, Scharfman HE, Lavenex P, The dentate gyrus: fundamental neuroanatomical organization (dentate gyrus for dummies), Prog. Brain Res 163 (2007) 3–22. [PubMed: 17765709]

[36]. Amaral DG, Ishizuka N, Claiborne B, Chapter 1 Chapter Neurons, numbers and the hippocampal network, Progr. Brain Res (1990) 1–11, 10.1016/s0079-6123(08)61237-6.

[37]. Kempermann G, Kuhn HG, Gage FH, Genetic influence on neurogenesis in the dentate gyrus of adult mice, Proc. Natl. Acad. Sci. U. S. A 94 (1997) 10409–10414. [PubMed: 9294224]

[38]. Spalding KL, Bergmann O, Alkass K, Bernard S, Salehpour M, Huttner HB, Boström E, Westerlund I, Vial C, Buchholz BA, Possnert G, Mash DC, Druid H, Frisén J, Dynamics of hippocampal neurogenesis in adult humans, Cell 153 (2013) 1219–1227. [PubMed: 23746839]

[39]. Guzman SJ, Schlögl A, Frotscher M, Jonas P, Synaptic mechanisms of pattern completion in the hippocampal CA3 network, Science 353 (2016) 1117–1123. [PubMed: 27609885]

[40]. Majani E, Erlanson R, Yaser Abu-Mostafa S, On the k-winners-take-all network, (1989), pp. 634–642.

[41]. Glorot X, Bengio Y, Understanding the difficulty of training deep feedforward neural networks, Proc. Mach. Learn. Res 9 (2010) 249–256.

[42]. Rojas R, The Backpropagation Algorithm, Neural Netw. (1996) 149–182, 10.1007/978-3-642-61068-4_7.

[43]. Kee N, Teixeira CM, Wang AH, Frankland PW, Preferential incorporation of adult-generated granule cells into spatial memory networks in the dentate gyrus, Nat. Neurosci 10 (2007) 355–362. [PubMed: 17277773]

[44]. Pardi MB, Ogando MB, Schinder AF, Marin-Burgin A, Differential inhibition onto developing and mature granule cells generates high-frequency filters with variable gain, Elife 4 (2015) e08764. [PubMed: 26163657]

[45]. Buckmaster PS, Strowbridge BW, Kunkel DD, Schmiege DL, Schwartzkroin PA, Mossy cell axonal projections to the dentate gyrus molecular layer in the rat hippocampal slice, Hippocampus 2 (1992) 349–362. [PubMed: 1284975]

[46]. Chawla MK, Guzowski JF, Ramirez-Amaya V, Lipa P, Hoffman KL, Marriott LK, Worley PF, McNaughton BL, Barnes CA, Sparse, environmentally selective expression of Arc RNA in the upper blade of the rodent fascia dentata by brief spatial experience, Hippocampus 15 (2005) 579–586. [PubMed: 15920719]

[47]. Toni N, Laplagne DA, Zhao C, Lombardi G, Ribak CE, Gage FH, Schinder AF, Neurons born in the adult dentate gyrus form functional synapses with target cells, Nat. Neurosci 11 (2008) 901–907. [PubMed: 18622400]

[48]. Richards BA, Frankland PW, The persistence and transience of memory, Neuron 94 (2017) 1071–1084, 10.1016/j.neuron.2017.04.037. [PubMed: 28641107]

[49]. Imayoshi I, Sakamoto M, Ohtsuka T, Takao K, Miyakawa T, Yamaguchi M, Mori K, Ikeda T, Itohara S, Kageyama R, Roles of continuous neurogenesis in the structural and functional integrity of the adult forebrain, Nat. Neurosci 11 (2008) 1153–1161. [PubMed: 18758458]

[50]. Cahill SP, Yu RQ, Green D, Todorova EV, Snyder JS, Early survival and delayed death of developmentally-born dentate gyrus neurons, Hippocampus 27 (2017) 1155–1167. [PubMed: 28686814]

[51]. Deisseroth K, Singla S, Toda H, Monje M, Palmer TD, Malenka RC, Excitation-neurogenesis coupling in adult neural stem/progenitor cells, Neuron 42 (2004) 535–552. [PubMed: 15157417]

[52]. Ishikawa R, Fukushima H, Frankland PW, Kida S, Hippocampal neurogenesis enhancers promote forgetting of remote fear memory after hippocampal reactivation by retrieval, Elife 5 (2016), 10.7554/eLife.17464.

[53]. Gao A, Xia F, Guskjolen AJ, Ramsaran AI, Santoro A, Josselyn SA, Frankland PW, Elevation of hippocampal neurogenesis induces a temporally graded pattern of forgetting of contextual fear memories, J. Neurosci 38 (2018) 3190–3198. [PubMed: 29453206]

[54]. Cuartero MI, de la Parra J, Pérez-Ruiz A, Bravo-Ferrer I, Durán-Laforet V, García-Culebras A, García-Segura JM, Dhaliwal J, Frankland PW, Lizasoain I, Moro MÁ, Abolition of aberrant neurogenesis ameliorates cognitive impairment after stroke in mice, J. Clin. Invest 129 (2019) 1536–1550. [PubMed: 30676325]

[55]. Brea J, Urbanczik R, Senn W, A normative theory of forgetting: lessons from the fruit fly, PLoS Comput. Biol 10 (2014) e1003640. [PubMed: 24901935]

[56]. Appleby PA, Wiskott L, Additive neurogenesis as a strategy for avoiding interference in a sparsely-coding dentate gyrus, Network 20 (2009) 137–161. [PubMed: 19731146]

[57]. Cameron HA, McKay RD, Adult neurogenesis produces a large pool of new granule cells in the dentate gyrus, J. Comp. Neurol 435 (2001) 406–417. [PubMed: 11406822]

[58]. Rolls ET, A quantitative theory of the functions of the hippocampal CA3 network in memory, Front. Cell. Neurosci 7 (2013), 10.3389/fncel.2013.00098.

[59]. Rolls E, Treves A, Neural Networks and Brain Function, (1997), 10.1093/acprof:oso/9780198524328.001.0001.

[60]. Ge S, Yang C-H, Hsu K-S, Ming G-L, Song H, A critical period for enhanced synaptic plasticity in newly generated neurons of the adult brain, Neuron 54 (2007) 559–566. [PubMed: 17521569]

[61]. Li L, Sultan S, Heigele S, Schmidt-Salzmann C, Toni N, Bischofberger J, Silent synapses generate sparse and orthogonal action potential firing in adult-born hippocampal granule cells, Elife 6 (2017), 10.7554/eLife.23612.

[62]. Morris RGM, Moser EI, Riedel G, Martin SJ, Sandin J, Day M, O'Carroll C, Elements of a neurobiological theory of the hippocampus: the role of activity-dependent synaptic plasticity in memory, Philos. Trans. R. Soc. Lond. B Biol. Sci 358 (2003) 773–786. [PubMed: 12744273]

[63]. Morris RGM, Garrud P, Rawlins JNP, O'Keefe J, Place navigation impaired in rats with hippocampal lesions, Nature 297 (1982) 681–683, 10.1038/297681a0. [PubMed: 7088155]

[64]. Packard MG, Hirsh R, White NM, Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems, J. Neurosci 9 (1989) 1465–1472. [PubMed: 2723738]

[65]. Paylor R, Zhao Y, Libbey M, Westphal H, Crawley JN, Learning impairments and motor dysfunctions in adult Lhx5-deficient mice displaying hippocampal disorganization, Physiol. Behav 73 (2001) 781–792. [PubMed: 11566211]

[66]. Deacon RMJ, Rawlins JNP, Learning impairments of hippocampal-lesioned mice in a paddling pool, Behav. Neurosci 116 (2002) 472–478, 10.1037/0735-7044.116.3.472. [PubMed: 12049328]

[67]. Aronov D, Nevers R, Tank DW, Mapping of a non-spatial dimension by the hippocampal-entorhinal circuit, Nature 543 (2017) 719–722. [PubMed: 28358077]

[68]. Kim J, Castro L, Wasserman EA, Freeman JH, Dorsal hippocampus is necessary for visual categorization in rats, Hippocampus 28 (2018) 392–405. [PubMed: 29473984]

[69]. Rajji T, Chapman D, Eichenbaum H, Greene R, The role of CA3 hippocampal NMDA receptors in paired associate learning, J. Neurosci 26 (2006) 908–915. [PubMed: 16421310]

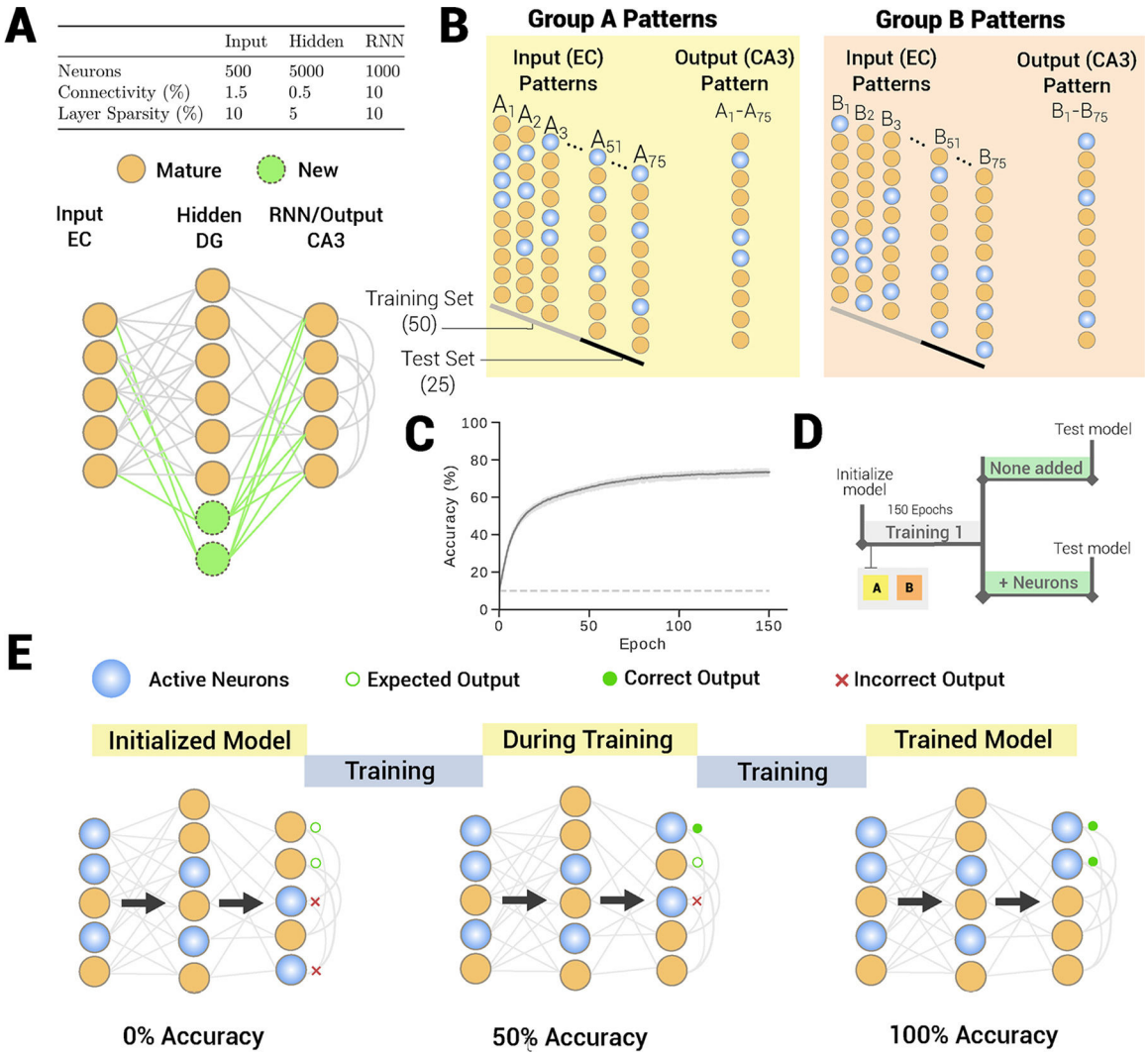

Fig. 1.

Implementing neurogenesis in an ANN. **(A)** An illustration of our neurogenesis model and general properties of model. The three layers represent the EC, DG and CA3 in sequential order. Mature (orange) and new (green) DG hidden neurons receive inputs from the EC and send projections to the CA3 in a sparse manner. Furthermore, the CA3 is modelled as an RNN which has recurrent connections to itself. (B) The categorization task involves generating A and B input patterns that draw from subsets of the neurons arbitrarily assigned to A or B output patterns. These patterns are split into training (50) and test (25) sets. (C) The learning curve of the network on the AB categorization task. The network is able to produce the correct output patterns (~70% solid line) well above chance levels (~10%, dashed line). This is the mean learning curve across 20 initializations of the model. (D) An illustration of the experimental design. Networks were first trained on AB categorization, followed by the addition of new neurons to the DG (neurogenesis). Models with and without neurogenesis were then tested on the AB categorization. (E) An illustration of how the network learns and performance is measured. As the network is presented input patterns,

we measure the Hamming distance between the actual and expected outputs and convert the measure to an accuracy.

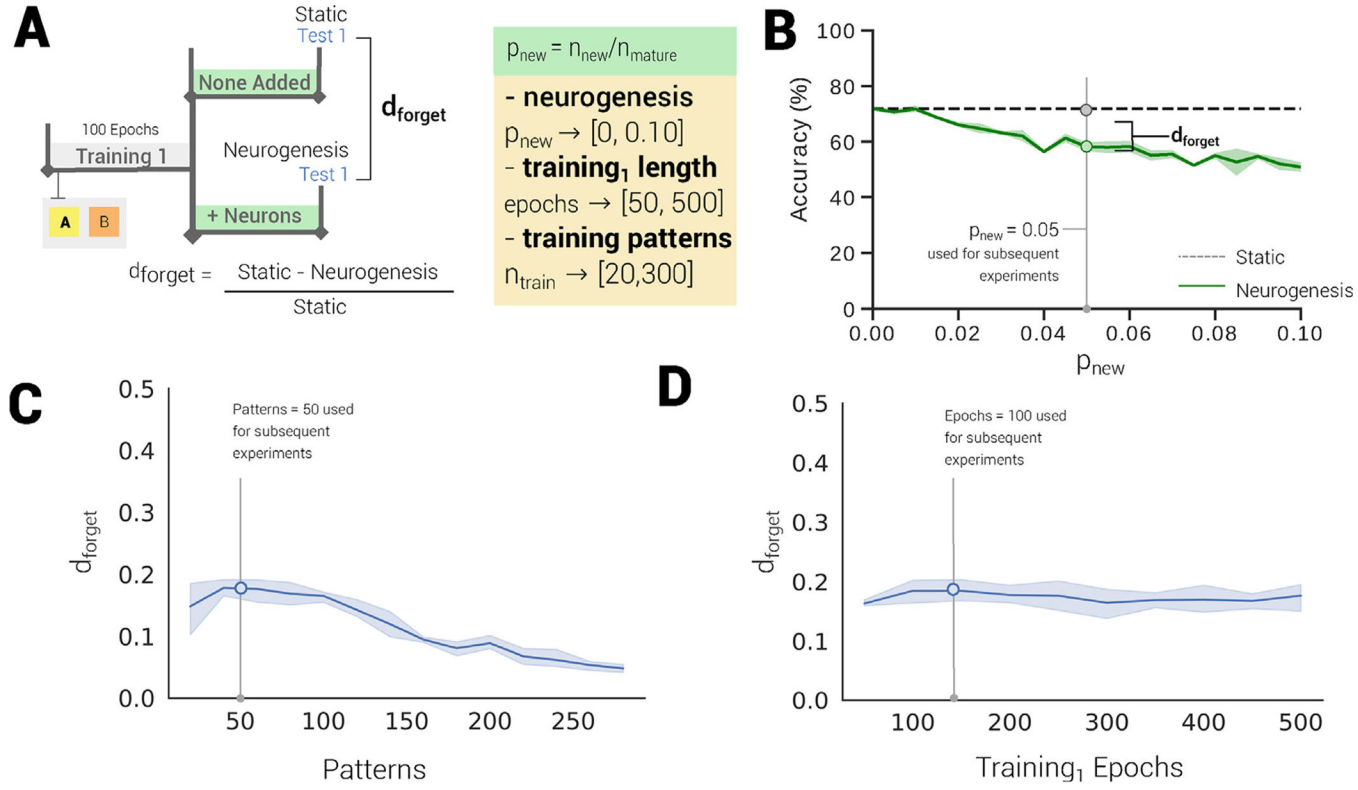

Fig. 2.

Forgetting impairs retrieval of categorization memory. (A) Illustration of forgetting experiment.(Left) Networks were first trained on AB categorization, followed by the addition of new neurons to the DG (neurogenesis). Models with (neurogenesis) and without neurogenesis (baseline) were then tested on the AB categorization. The difference in performance between baseline and neurogenesis is captured by the variable d_{forget}. (Right) We vary p_{new} between 0 and 0.10, which is the amount of new neurons added as a proportion of the DG size (5000). (B) Plot of test error as we vary p_{new}. Networks with neurogenesis have worse performance as more neurons are added. We use a value of p_{new} = 0.05 in all subsequent experiments unless otherwise indicated. (C) Plot of mean d_{forget} with overtraining on increasing training data (solid line). (D) Plot of mean d_{forget} with overtraining for more epochs (solid line).

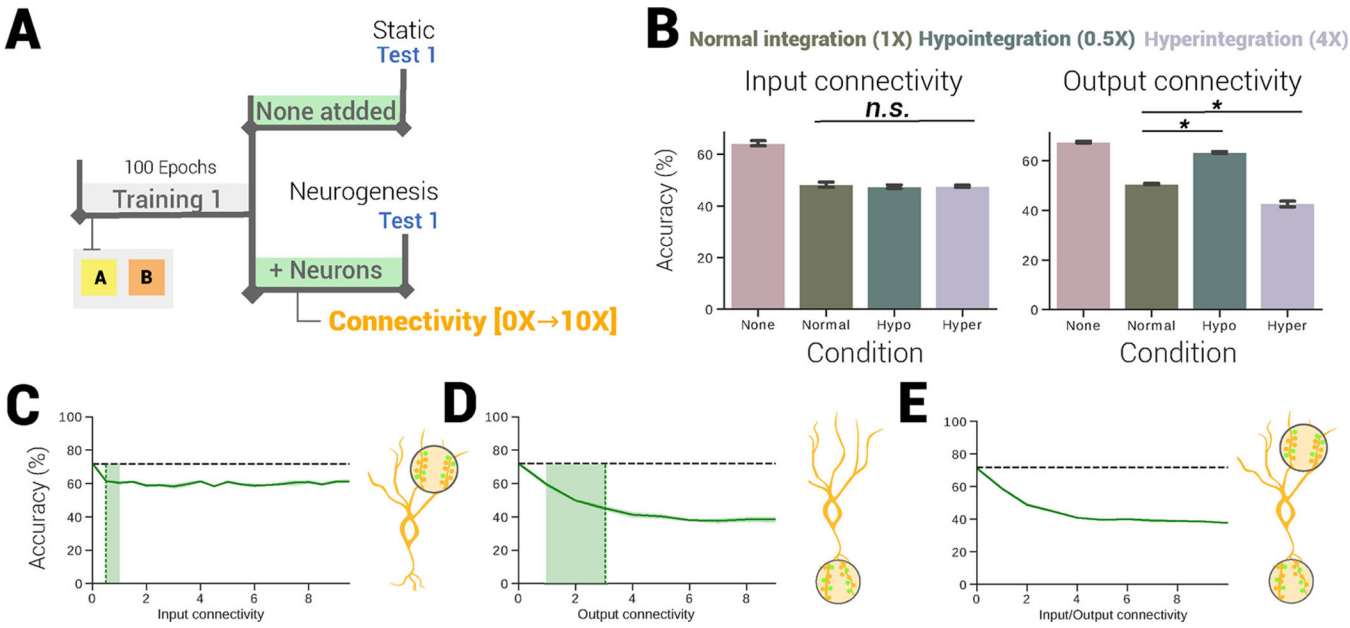

Fig. 3.

The impact of input and output connectivity on forgetting. (A) Illustration of forgetting experiment. Networks were first trained on AB categorization, followed by the addition of new neurons to the DG (neurogenesis). Models with neurogenesis and without neurogenesis (static) were then tested on the AB categorization. We varied the connectivity of the hidden layer's input and output connections between 0 to 10X (relative to input/output connectivity of the mature neurons). (B) Plot of the varying effects of normal, hypo and hyper-integration of input and output connections vs. static (i.e., no neurogenesis or none). Left: Input connectivity significantly decreased AB categorization (Accuracy × Condition [$F(1, 10) = 88.65$, $p = 3.50 \times 10^{10}$]) where high (hyper), low (hypo) and normal levels of input connectivity were not significantly different. Right: There was a significant effect of output connectivity on AB categorization (Accuracy × Condition [$F(1, 10) = 200.10$, $p = 2.17 \times 10^{06}$]) whereby hyperintegration of output connectivity increased forgetting ($p < 0.001$, Tukey HSD post-hoc test), and hypointegration decreased forgetting ($p < 0.001$, Tukey HSD post-hoc test) relative to the normal levels of output connectivity. (C-E) Plots of performance at varying levels of input (C), output (D) and concurrent input/output (E) connectivity. Low rates of input connectivity were sufficient to cause forgetting. However, increasing levels of input connectivity did not produce further forgetting. Low rates of output connectivity were sufficient to cause forgetting, and forgetting increased with increased output connectivity. Similarly, forgetting increased when both input and output connectivity were varied concurrently.

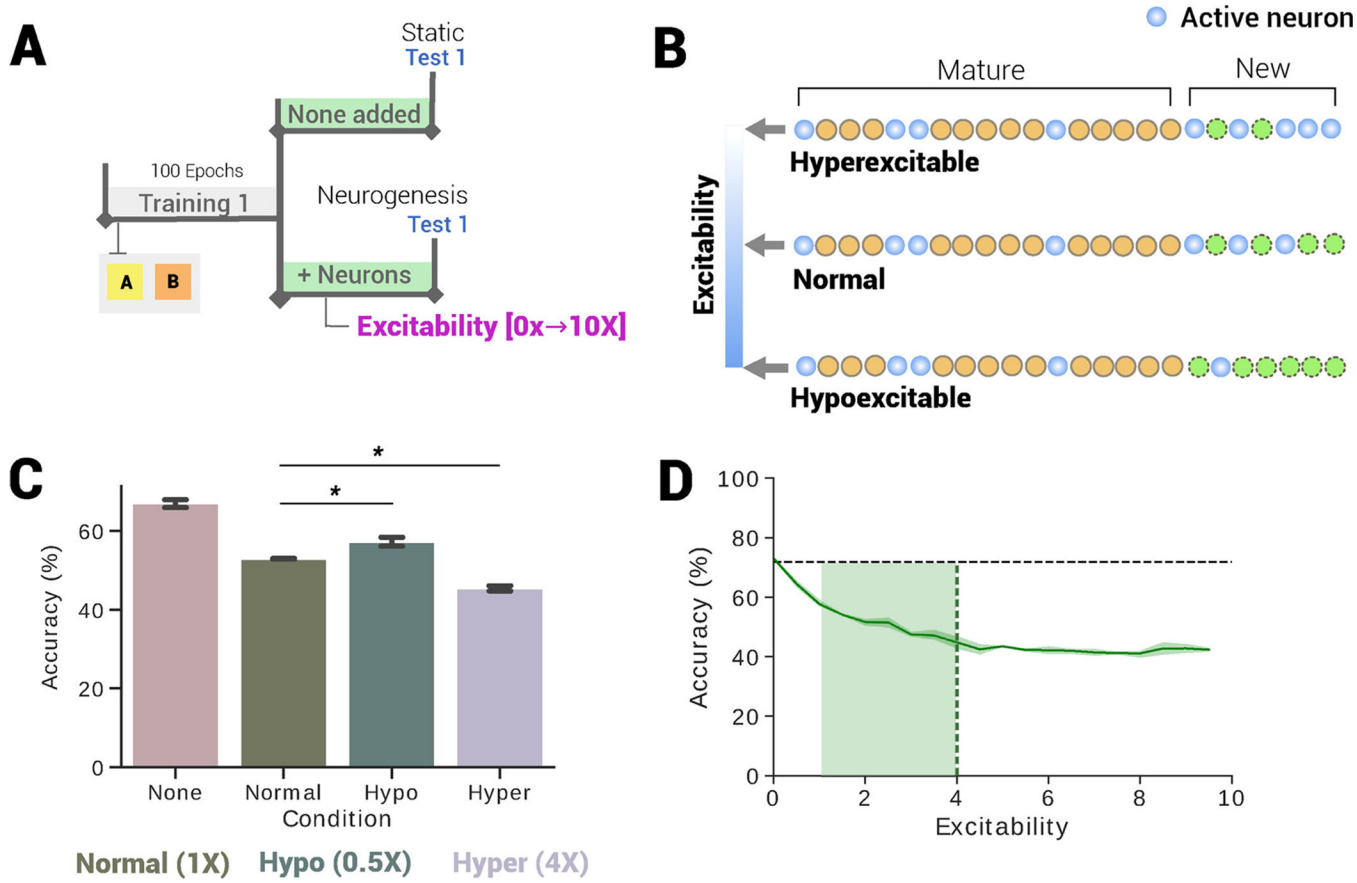

Fig. 4.

The impact of neuronal excitability on forgetting. (A) Illustration of forgetting experiment. We varied the excitability of the pool of immature neurons between 0 to 10X that of the mature neurons. (B) Illustration of excitability in our model. Neurons are constrained by a KWTA rule such that only a sparse set of neurons (0.05 of the DG) with the highest activities are able to pass on information to the next layer. New neurons can have this sparsity reduced (hyperexcitability) or increased (hypoexcitability) beyond 0.05. (C) Plot of the varying effects of normal, hypo- and hyperexcitability of new neurons. Excitability significantly decreased AB categorization in the neurogenesis group (Accuracy × Excitability [$F(1, 10) = 220.76$, $p = 4.97 \times 10^{08}$]) whereby hyperexcitability increased forgetting ($p < 0.001$, Tukey HSD post-hoc test), and hypoexcitability decreased forgetting ($p < 0.001$, Tukey HSD post-hoc test) relative to the normal levels of excitability. (D) Low rates of excitability were sufficient to cause forgetting, and forgetting increased with reduced sparsity of the new neuron population.

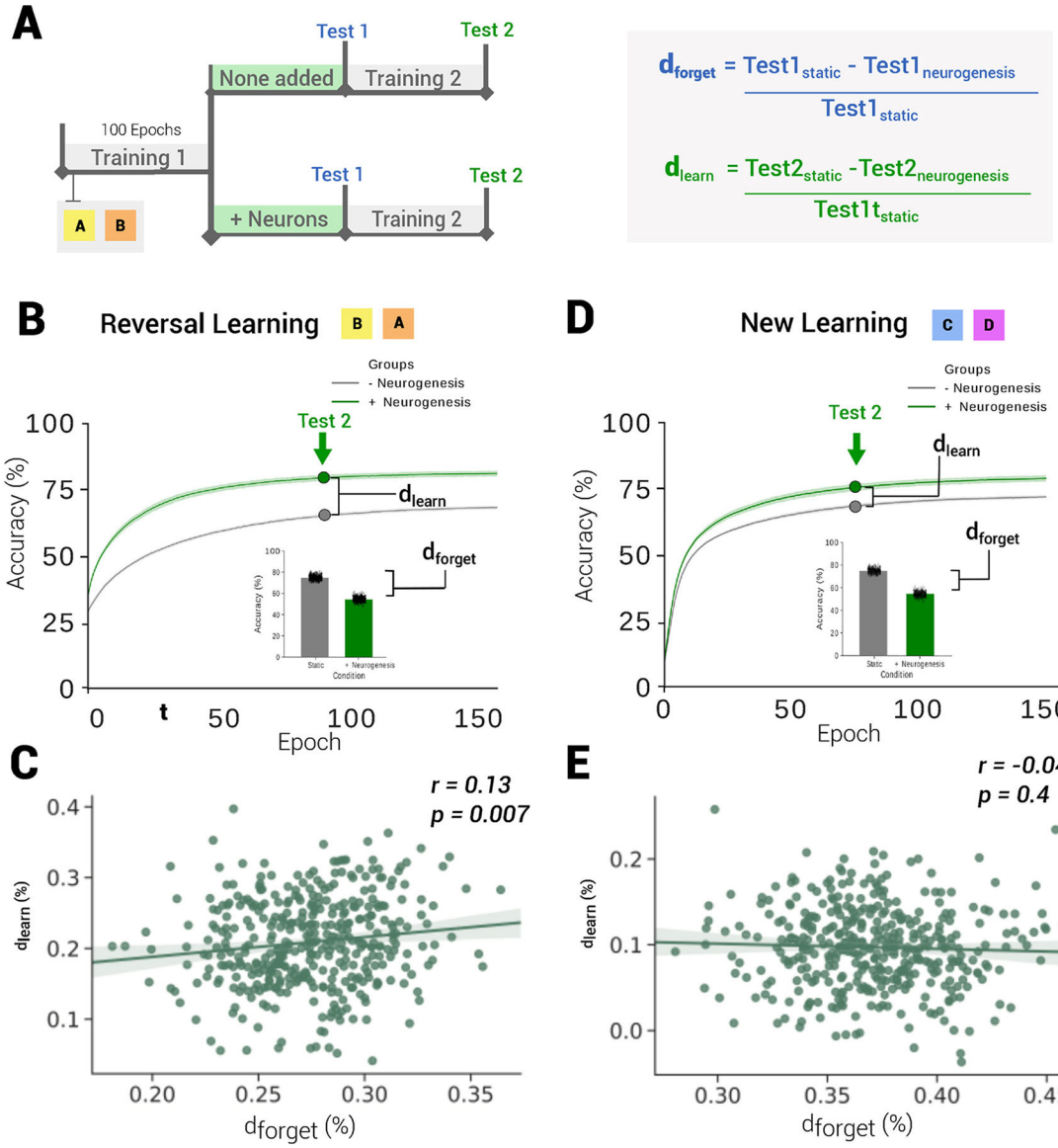

Fig. 5.

Forgetting moderately predicts reversal learning enhancement. (A) Illustration of the experiment (left). d_{forget} and d_{learn} calculated from Test 1 and Test 2, respectively, between neurogenesis and static groups (right) (B). Plot of Training 2 of reversal learning curves for 250 initialized neurogenic and static models. (Inset) Plot of the performance of the same neurogenic networks after adding neurons, compared to static networks. (C) Plot of d_{learn} vs. d_{forget} of the 250 simulations of reversal learning on BA categorization, fit with a linear regression ($r = 0.133$, $p = 0.007$). (D). Plot of Training 2 of new learning curves for 250 initialized neurogenic and static models. (Inset) Plot of the performance of the same neurogenic networks after adding neurons, compared to static networks. (E) Plot of d_{learn} vs. d_{forget} of the 250 simulations of new learning on C, D categorization, fit with a linear regression ($r = -0.04$, $p = 0.4$).

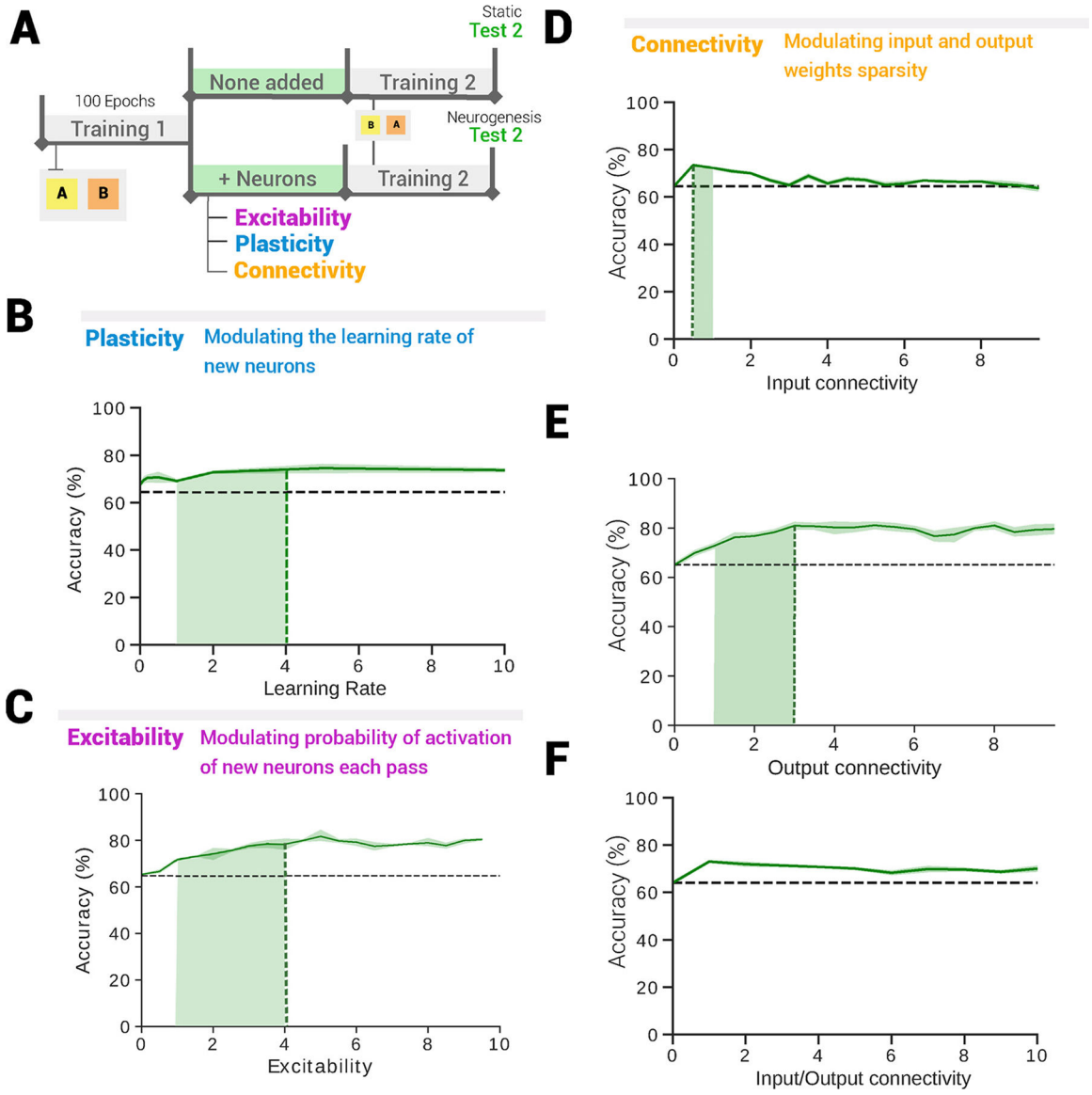

Fig. 6.

Exploring parameters involved in neurogenesis-mediated enhancement in reversal learning. (A) Illustration of the experiment. Excitability, plasticity, and input and output connectivity of new neurons are varied independently. (B) Plot comparing the performance of static and neurogenesis networks when varying plasticity (Group × Learning Rate [$F(2, 20) = 0.51$, $p = 0.477$]). (C) Plot comparing the performance of static and neurogenesis networks when varying excitability (Group × Excitability [$F(2, 20) = 179.48$, $p = 3.43 \times 10^{30}$]). (D) Plot comparing performance of static and neurogenesis networks when varying input connectivity (Group × Input connectivity [$F(2, 20) = 32.98$, $p = 5.46 \times 10^{86}$]). (E) Plot comparing performance of static and neurogenesis networks when varying output connectivity (Group × Output connectivity [$F(2, 20) = 199.34$, $p = 1.10 \times 10^{31}$]). (F) Plot comparing performance of static and neurogenesis networks when varying both input and output connectivity (Group × Both Input/Output Connectivity [$F(2, 20) = 0.25$, $p = 0.617$]);

Fig. 6F). All parameters of new neurons are shown as a relative value to values used for mature neurons. All plots show the mean and standard error of the mean of reversal learning across 20 experiments for each parameter. Horizontal dashed lines indicate performance of static networks. Shaded regions indicate the parameter range from values used in mature neurons to those used in immature neurons for figures 5,7,8 (vertical dashed lines).

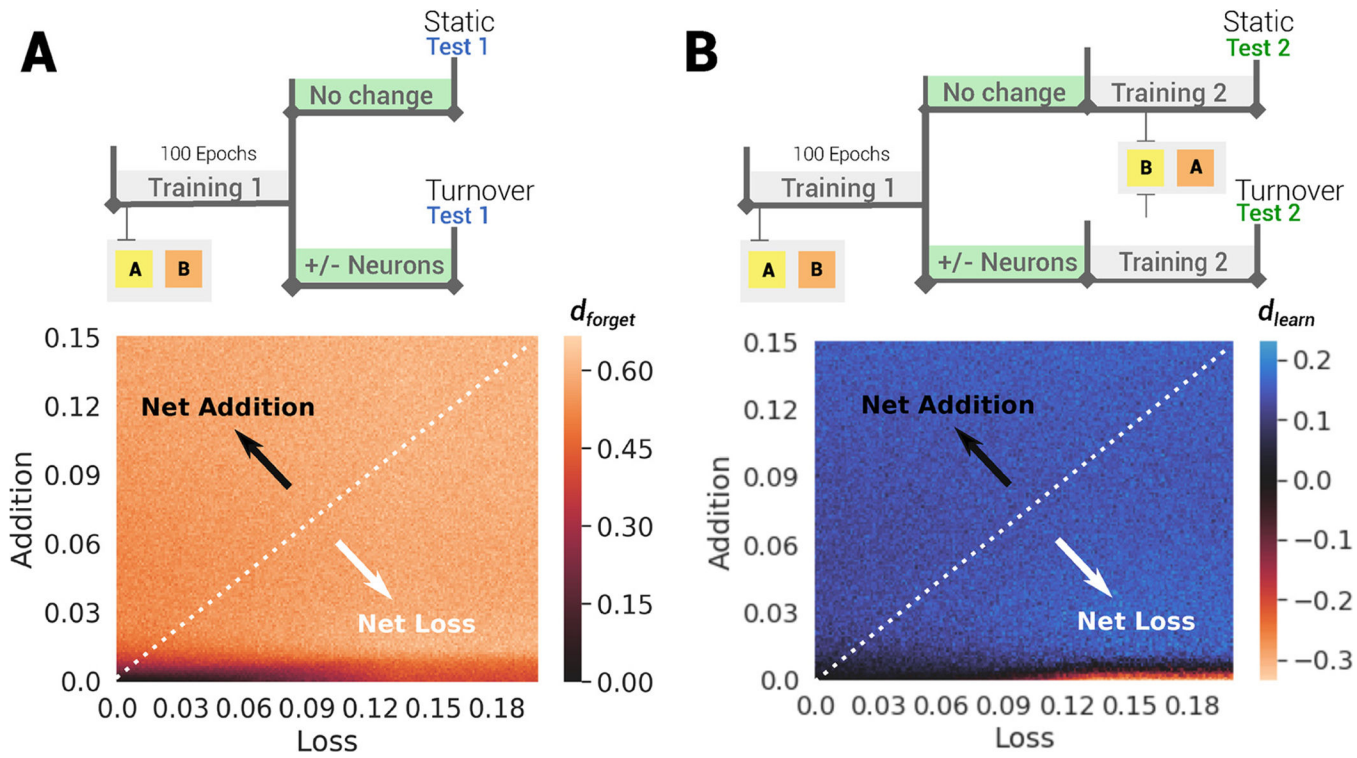

Fig. 7.

Addition drives forgetting and reversal learning enhancement. (A) Heatmap of the d_{forget} across varying fractions of loss and addition. Diagonal indicates balanced loss and addition. (B) Heatmap of d_{learn} across varying fractions of loss and addition. Each cell represents the mean of 20 experiments with different model initializations.

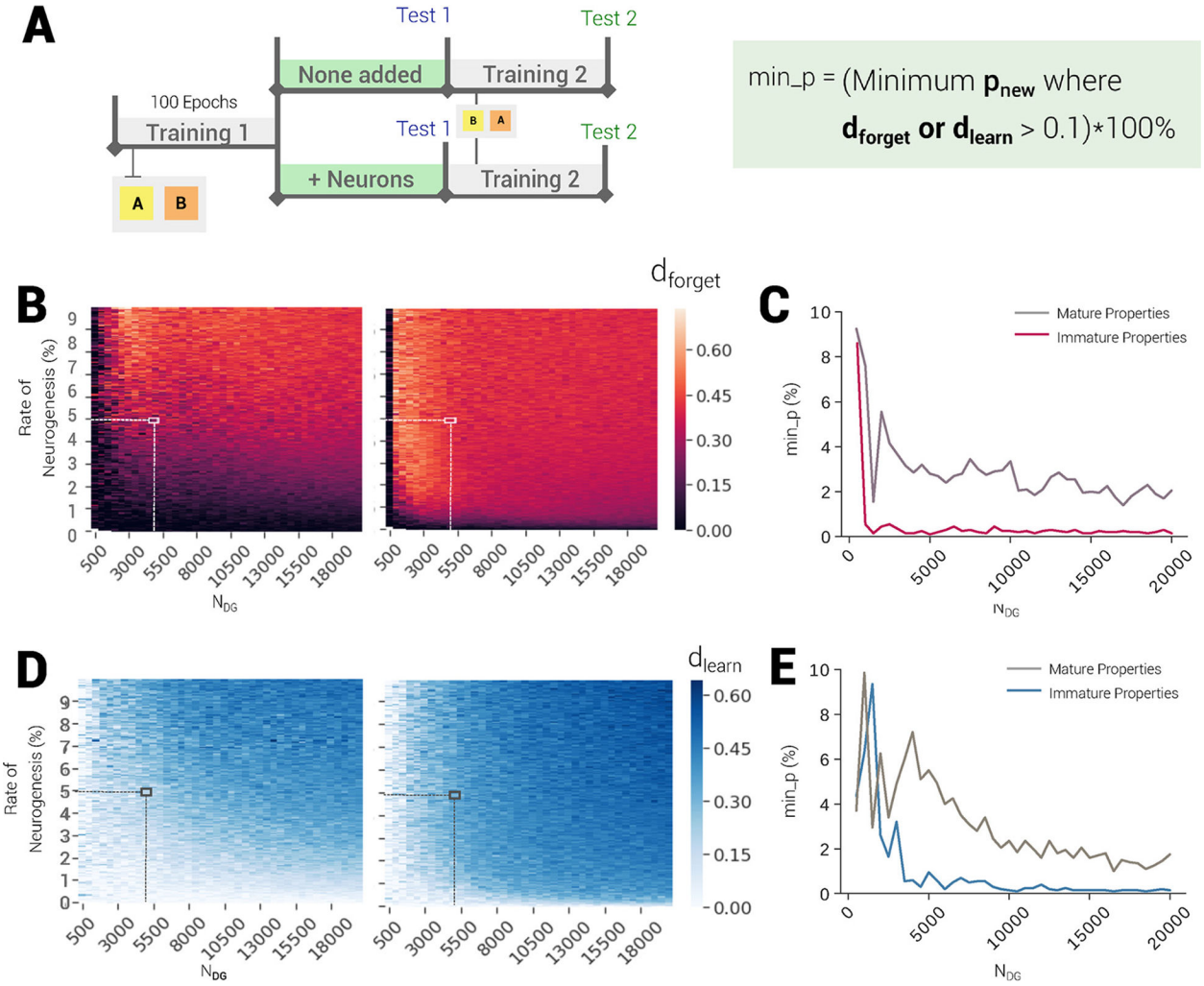

Fig. 8.

Forgetting and enhanced reversal learning at varying rates of neurogenesis and model size. (A) Illustration of experiment. (B) Left: Heatmap of d_{forget} across different rates of neurogenesis and network sizes using mature neuron properties in new neurons. Right: Heatmap of d_{forget} across different rates of neurogenesis and network sizes using biologically-inspired properties of new neurons. Each cell is the mean d_{forget} of 5 experiments with different model initializations. (C) Plot of the minimum p_{new} where a d_{forget} is observed in **B** at different network sizes when using the mature neuron properties (gray) or new neuron properties (red). (D) Heatmap of d_{learn} across different rates of neurogenesis and network sizes. Each cell is the mean d_{learn} of 5 experiments with different model initializations. (E) Plot of the minimum p_{new} where a d_{learn} of 0.10 is observed in **D** at different network sizes when using the mature neuron properties (gray) or new neuron properties (blue).