RESEARCH ARTICLE

Thunderstruck: The ACDC model of flexible sequences and rhythms in recurrent neural circuits

Cristian Buc Calderon [1,2,3] *, Tom Verguts [2], Michael J. Frank [1,3]

1 Department of Cognitive, Linguistic & Psychological Sciences, Brown University, Providence, Rhode Island, United States of America, **2** Department of Experimental Psychology, Ghent University, Ghent, Belgium, **3** Carney Institute for Brain Science, Brown University, Providence, Rhode Island, United States of America

* cbuccald@gmail.com

Abstract

Adaptive sequential behavior is a hallmark of human cognition. In particular, humans can learn to produce precise spatiotemporal sequences given a certain context. For instance, musicians can not only reproduce learned action sequences in a context-dependent manner, they can also quickly and flexibly reapply them in any desired tempo or rhythm without overwriting previous learning. Existing neural network models fail to account for these properties. We argue that this limitation emerges from the fact that sequence information (i.e., the position of the action) and timing (i.e., the moment of response execution) are typically stored in the same neural network weights. Here, we augment a biologically plausible recurrent neural network of cortical dynamics to include a basal ganglia-thalamic module which uses reinforcement learning to dynamically modulate action. This "associative cluster-dependent chain" (ACDC) model modularly stores sequence and timing information in distinct loci of the network. This feature increases computational power and allows ACDC to display a wide range of temporal properties (e.g., multiple sequences, temporal shifting, rescaling, and compositionality), while still accounting for several behavioral and neurophysiological empirical observations. Finally, we apply this ACDC network to show how it can learn the famous "Thunderstruck" song intro and then flexibly play it in a "bossa nova" rhythm without further training.

Author summary

How do humans flexibly adapt action sequences? For instance, musicians can learn a song and quickly speed up or slow down the tempo, or even play the song following a completely different rhythm (e.g., a rock song using a bossa nova rhythm). In this work, we build a biologically plausible network of cortico-basal ganglia interactions that explains how this temporal flexibility may emerge in the brain. Crucially, our model factorizes sequence order and action timing, respectively represented in cortical and basal ganglia dynamics. This factorization allows full temporal flexibility, i.e. the timing of a learned

action sequence can be recomposed without interfering with the order of the sequence. As such, our model is capable of learning asynchronous action sequences, and flexibly shift, rescale, and recompose them, while accounting for biological data.

Introduction

Learning and manipulating sequential patterns of motor output are essential for virtually all domains of human behavior. For instance, musicians can learn multiple precise spatiotemporal sequences, each with their own rhythm. They can later modify the tempo to a learned sequence, or even apply a completely different rhythm, e.g., perform a rock song with a bossa nova rhythm. Thus, musicians can quickly and flexibly manipulate action timing in action sequences. Similar capabilities abound in other domains, such as language production and athletics.

Precisely timed action sequences are thought to emerge from dynamical neural patterns of activity. In particular, sparse sequential activity patterns observed in basal ganglia [1–5], hippocampus [6–9] and the cortex [10–12] are thought to provide a temporal (ordinal) signal for these action sequences to emerge. However, although seminal modeling work has been carried out to understand how sequences emerge in neural networks [13–15], the mechanistic and dynamic principles by which these neural patterns afford sequential flexibility remain unknown. While several neural network models of corticostriatal circuits exist, these are typically applied to single shot stimulus-action pairings rather than sequential choices, despite extensive evidence that basal ganglia is implicated in such sequential behaviors [16,17] (but see [18] for a nuanced view).

We sought to develop a biologically plausible neural computational model of cortico-basal ganglia circuitry sufficiently powerful to learn arbitrary sequences (e.g., scales) and easily adjust their timing and expression on the fly. In particular, we aimed for the network to be able to learn multiple arbitrary sequences and to allow for temporal asynchrony, shifting, rescaling, and compositionality. We define these terms more precisely below.

Neurocomputational models of sequence production can be broadly categorized in three classes, each with their advantages and disadvantages in computational power and their ability to account for behavioral and neural data.

- In <u>associative chain models</u> (also termed synfire chain [19–21]), activation flows sequentially from one neuron (or neuronal population) to another through feedforward connections [22]. The sequence emerges from the hard-wired structure of the chain. Associative chain models produce sequential and persistent neural activity, as observed empirically [23,24]. They can deal with inherent compression of sequential activity, and learn to produce each action in the sequence at any desired precise time [22]. However, these models are not equipped to facilitate *temporal rescaling*: the finding that learned action sequences can be sped up (compressed) or slowed down (dilated) without the need to overwrite previous learning [25,26]. Moreover, it is unclear how these models implement *temporal shifting*: the ability to start the action sequence earlier or later in time, without modifying the action sequence structure. Chain models also do not straightforwardly allow networks to encode more than a single sequence, given their hard-wired nature.

- <u>Cluster-based models</u> also involve a chained sequence of activation, but this sequence is learned via cell assemblies (clusters) that are chained within a recurrent neural network (RNN), for instance, through spike timing dependent plasticity [27–32]. Hence, the chaining

emerges from synaptic learning rather than being hard-wired. Once the chain is learnt, a single initial input pulse to the RNN induces a sequential activation whereby activation flows from one cluster to another. Cluster-based models allow temporal rescaling [28] while also producing sequential and persistent patterns of activity [27]. Furthermore, they provide a simple mechanism allowing a network to encode multiple sequences. By selectively activating a specific cluster within the RNN, only the cluster "in line" (i.e., connected to the previous cluster) is activated sequentially. Therefore, the RNN can encode multiple sequential actions by learning (and selectively activating) distinct cluster chains encoded in the connectivity matrix [28]. Yet, it is unclear how these models could facilitate action sequences with *temporal asynchrony*: the ability to learn, and flexibly manipulate, motor sequences with varying inter-action intervals (an advantage of associative chain models [22]). Indeed, cluster-based models can flexibly manipulate sequences; however, these sequences are typically iso-synchronous [28]. In addition, although emergent connectivity within and between clusters (or units) can arise via unsupervised learning [33], this connectivity crucially depends on the sequential nature and timing of the teaching input signal to the distinct subsets of the RNN.

- State-space models [34] do not assume a chaining structure at all. Based on a (sparse) randomly connected RNN structure, these models produce a neural trajectory that evolves in high-dimensional space which can be used as a temporal basis to perform a range of complex tasks [35]. However, to reliably reproduce the same task, neural trajectories need to be robust to noise. To that end, state-space models typically harness a noiseless neural trajectory (based on any random connectivity) which is then subsequently used as a continuous teaching signal in presence of noise [34,36,37]. Alternatively, each individual unit in the RNN can be taught via supervised learning to reproduce the neural activity of an empirical dataset [38]. The resultant learned neural trajectory [39] acts as a robust travelling wave that can be decoded by downstream neurons to produce highly complex and flexible motor sequences [40]. However, to reproduce reliable motor sequences, state-space models require highly supervised teaching signals specifying the full neural trajectory and non-biological learning mechanisms (e.g., residual least squares learning algorithms [38,41]). Recent work has shown that biological learning rules using local information can effectively learn complex (sequential) tasks [42–44] (albeit not as effectively as non-biological rules). State-space models can also implement a rudimentary form of temporal rescaling, in that they can rescale the timing of the execution of a single motor response [45], and iso-synchronous action sequences (e.g., index tapping at a steady rhythm) [34]. However, these models do not support temporal rescaling in the more general case (i.e., asynchronous action sequences). Furthermore, given their focus on cortical networks, these models do not address the growing evidence that action sequences unfold over multiple levels within cortico-basal ganglia-thalamic loops, with attractor state switches occurring in the prefrontal cortex [46] and action timing represented in the basal ganglia [2,4].

- Finally, none of the models have tackled how a learned sequence at a particular tempo can be executed with a completely different tempo which may have been learned for a different sequence (e.g., applying a bossa nova rhythm to a rock song). We refer to this ability as *temporal compositionality*.

In sum, all models can account for distinct functionalities in sequence production, but fail to provide a plausible neurocomputational mechanism from which most fundamental abilities–temporal asynchrony, shifting, rescaling, compositionality–can emerge and interact. These limitations arise from a property common to all action sequence models: action identity,

timing and sequence order are represented jointly within the recurrent weights of the network. In the reinforcement learning domain, such joint coding of task features facilitates only rigid forms of generalization and transfer, whereas the ability to code task features compositionally facilitates more robust transfer [47] that can better account for human behavior [48]. However, the mechanisms for such compositionality in neural networks remains unknown.

Here, we develop a biologically inspired (we further discuss biological plausibility of our model in the Discussion) RNN called the associative cluster-dependent chain (ACDC) model. By combining strengths of the associative chain and cluster-based models, ACDC accounts for biological data. As we show below, the novelty in our model is twofold. First, we propose a biologically-plausible model of cortico-basal ganglia-thalamic loops that decomposes the functions of cortex and basal ganglia and learns sequences based on simple local and (biologically motivated) supervised learning rules. Second, this decomposition affords greater flexibility in generating desired action sequences, supporting temporal asynchrony, shifting, rescaling, and compositionality in a single model. Crucially, our model factorizes action sequence features within the circuit, with cortical RNN representing latent states within a sequence, and BG controlling both the timing of the transitions from one state to the next and which actions are linked to sequence positions. Factorizing order and timing information by storing them separately in a premotor cortical RNN, which is dynamically gated by a basal ganglia-thalamus module, affords independent (and flexible) manipulation of sequence order and action timing, and thus increases computational flexibility.

Results

We start by providing the reader with an intuitive functioning of the ACDC model (Fig 1). The Methods section provides detailed mathematical formulation and further grounds the model within the context of neurophysiological observations on the premotor cortex (PMC) and the BG. The ACDC model comprises a context module encoding the sequences to be executed (e.g., which song is to be played), and is provided as input to a RNN. This input targets a subset of RNN excitatory units, which cluster together via Hebbian learning, encoding the first latent state in the sequence (but not its specific action). In turn, the G (for Go) units in the BG learn (also via Hebbian learning) to link this RNN cluster to the appropriate action (blue arrow 1 in Fig 1), allowing it to accumulate evidence for the first action in the sequence. The G node, part of a G-A-N triplet, projects excitatory connections to its correspondent A (for Action) node (blue arrow 2 in Fig 1) which learns (via a delta rule) weight values for these projections to fine-tune the appropriate timing for this particular action. The A node represents motor thalamus, and its activation has two important consequences. First, it sends a thalamostriatal back-projection to excite the N node (blue arrow 3C in Fig 1), which finally inhibits the G node via lateral connections from D2 to D1 medium spiny neurons [49]. Second, the thalamic A node triggers a transition in the RNN, via a combination of excitatory projections to another RNN cluster (blue arrow 3A in Fig 1), and to a shared inhibitory neuron (blue arrow 3B in Fig 1), consistent with evidence that thalamic units target both cortical excitatory and inhibitory neurons [50–52]. Thus, whenever an action is executed, the ratio of excitatory to inhibitory inputs to the RNN is perturbed in a way that induces a transition from the current cluster to the next cluster in line (targeted by the feedback projections of the current A node, blue arrow 3A in Fig 1) to be expressed (see Methods for more details).

Learning takes place over fast and slow time scales. Hebbian learning is fast and unfolds within the dynamics of a trial (i.e., during the evolution of an action sequence). In contrast, the delta rule is slow and is implemented between trials, via a signed error computed through the discrepancy between the action timing provided by the tutor and the generated action. Action
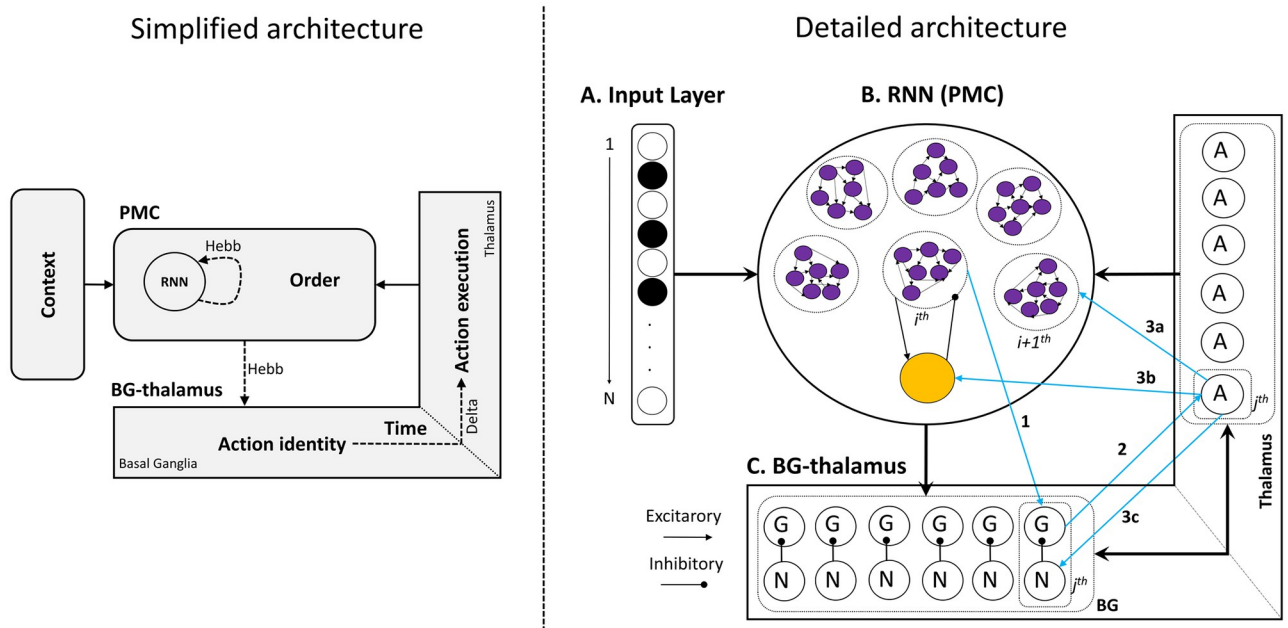
Fig 1. Simplified ACDC model architecture (left). An input context layer indicates which sequence needs to be learned or executed. The premotor cortex (PMC) is subtended by a RNN that learns (via Hebbian learning) to form clusters of excitatory neurons encoding order in the sequence, and which are regulated by an inhibitory neuron. In turn, each cluster learns to trigger action plans, topographically represented in the BG. Specific actions are executed in the thalamus at specific times based on learned connections from BG to thalamus. Motor activity is then fed back to the RNN, closing the cortico-basal ganglia loop. The unfolding of several iterations of this loop is responsible for the execution of precisely timed action sequences. Dashed lines are plastic connections, and the associated learning rule is indicated (Hebb = hebbian learning, Delta = delta rule). Note that the plastic connections from action identity to action execution correspond solely to the blue projection n°2 on the detailed architecture figure. **ACDC full model architecture (right). A. Input layer**: codes for contexts indicating the sequence to be learned/produced in a *N* length binary vector. **B. RNN**: represents recurrently interconnected neurons of the PMC, composed of a subset of interconnected neurons (i.e., clusters) that can give rise to sequential activation states after learning via cortico-basal ganglia loops. All excitatory nodes in the RNN project to a shared inhibitory neuron (orange node), which in turn inhibits all excitatory neurons (purple nodes; shown for just one cluster for visual simplicity). **C. The BG**: composed of two neuron types G (Go cells) and N (No Go cells). Go nodes accumulate evidence over time and excite Action (A) nodes in the BG output /thalamus layer. Once activity in a Go node reaches a specific threshold, the corresponding action is executed. Once executed, Action nodes reciprocally activate No Go nodes which in turn suppress Go nodes, shutting down action execution. **The thalamus**: is composed of Action nodes whose activity represents action execution. The jth Action node selectively projects excitatory connections to the $i+1^{th}$ cluster in the RNN, the shared inhibitory neuron and the jth No Go node in the BG. Light blue arrows represent the ith cortico-basal ganglia loop instance. The subindex i refers to the ordinal position in the sequence, i.e. the order the action possesses in the sequence. The subindex j represents the action that is associated to the ordinal position in the sequence.

sequences are learned sequentially: the model learns to produce the first action at the appropriate time, then the second, and so forth. Sequential learning improves motor execution [53–57], and is at the base of several theoretical models of motor sequence learning [58–60].

At a higher-level, order is encoded as a sequence of attractor states represented by persistent activation in distinct excitatory RNN unit clusters (cell assemblies). These clusters do not represent the actions themselves but rather their abstract order; the specific actions to be executed are learned via RNN projections to the BG and their timing is encoded in the weights of topographic projections to the motor thalamus. To optimize precise action timing, the weights between action identity (G unit activity) and execution (timing for a given action conditional on G unit activity) are learned via supervised learning (i.e., delta rule), perhaps summarizing the role of cerebellum in error corrective learning. This allows us to model tasks in which a tutor provides feedback (e.g., [61]; see Methods). Finally, feedback to the RNN from thalamic activity ultimately creates a cortico-basal ganglia loop. Each loop subtends the appropriate action order, identity and timing execution, allowing precisely timed action sequences to unfold. As we show below, our model architecture, allowing to uniquely encode timing

information in a distinct subset of the network (BG) than the one encoding order (PMC), will prove to display advantageous properties. In particular, being able to flexibly control, via external stimulation, the dynamics of the BG will result in a model displaying several temporal flexibility properties.

Learning precise spatiotemporal sequences

Simulation 1: Learning temporally asynchronous action sequences. Fig 2 shows simulation 1, where the ACDC model learns to produce a precisely timed, temporally asynchronous, action sequence (here, for 6 actions). The goal of the model is to produce each action sequentially at the appropriate time, here at 200, 250, 400, 700, 750 and 900 ms. This is an arbitrarily chosen timing sequence; the model can (learn to) produce any timed, synchronous (see below) or asynchronous, sequence. Fig 2A shows how the activity of each Action node progressively reaches the optimal time (depicted by color coded vertical dashed lines), reflected in a decrease in the action timing error (Fig 2B) and in the weight changes between Go and Action nodes (Fig 2C).

Fig 2D depicts the RNN connectivity matrix after learning (weights are zero before learning). Excitatory projections to the RNN from the input and motor layer are pseudo-random, with the restriction that two different projections never excite the same RNN neuron. These pseudo-random projections make it hard to visually identify the presence of clusters in Fig 2D; importantly however, this connectivity matrix does induce clustered dynamics (see S2 Video). Fig 2E shows how the ith cluster in the RNN learns to be (almost) selectively wired with the jth Go node. Finally, for completeness and transparency, Fig 2F portrays the dynamics of G (left) and N (right) nodes after learning of the sequence.

Temporal flexibility properties of the ACDC model

Having established learned RNN clusters during sequences, we now focus on the flexibility properties of the ACDC model after learning, without having to overwrite learned weights. First, we show that a previously learnt action sequence with *temporal asynchrony* can be flexibly reproduced. Second, we show that this sequence can be initiated earlier or later in time; we call this property *temporal shifting*. Third, we demonstrate how action sequences can be compressed or dilated, i.e., *temporal rescaling*. Fourth, we show how a given ordered sequence can be produced with a completely different tempo, a property that we refer to as *temporal compositionality*. Fifth, we describe how the model can also output sustained action execution. Finally, we show how the ACDC model can learn (a part of) the Thunderstruck song, which is then flexibly played on a bossa nova tempo; thereby recapitulating the temporal flexibility properties.

Simulation 2: Reproduction of previously learnt action sequence displaying temporal asynchrony. In simulation 1, we demonstrated that the ACDC model can learn precisely timed, temporally asynchronous, action sequences. In simulation 2, we freeze the weights and simply observe that the network can reproduce the sequence maintaining its precision in action timing (Fig 3A).

Simulation 3: Temporal shifting. The previous action sequence can be shifted in time, i.e., initiated earlier or later. Importantly, this shift can occur without changing the timing between actions (i.e., sequence timing is preserved). The ACDC model achieves flexible temporal shifting by implementing an additive positive (to start the sequence earlier) or negative (to start it later) input to the first Go node of the sequence, analogous to the top-down input from pre-SMA to striatum thought to bias starting points for evidence accumulation [62] (although similar effects could be implemented by dopaminergic modulation; see Discussion).
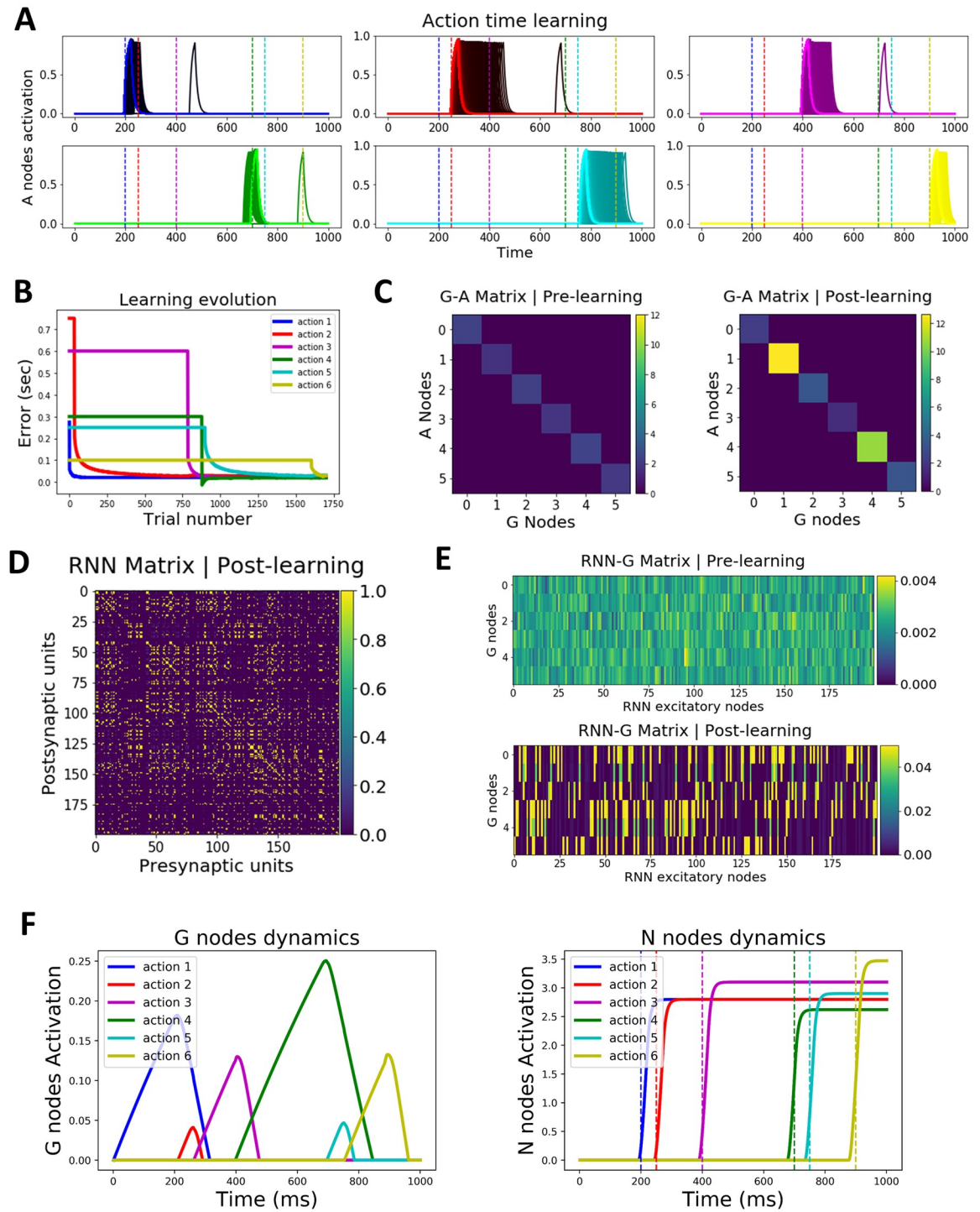
Fig 2. ACDC's learning dynamics. A. Learning a precisely timed action sequence. Each action execution (A node activation) is progressively shifted towards the optimal action time (depicted by the color coded vertical dashed line; x-axis represents time). Learning progresses from darker to brightest colors. The 'black trials' are early learning trials where the different shades are not distinguishable **B. Learning evolution**. Color coded traces represent the evolution of the error as a function of trial number for each action in the sequence. Learning unfolds sequentially, whereby timing errors are minimized for the first action before the second action starts learning. Therefore, each action (except action 1) starts off with a plateaued error level until the preceding action reaches the optimal time. Some action timings are learned faster than others because their optimal time weight value is closer to their initial value. The error is computed by subtracting the observed from the desired response time and plotted in seconds. **C. BG weights encode time**. Action timing is learned by changing the weights from BG Go nodes to thalamus Action nodes. The left and right panel show respectively the weights

values before and after learning. For instance, the second action (red trace in B) starts off being produced too slowly. Hence, weights increase until they produce the optimal action time for action 2. Color bars indicate weight values. **D. RNN connectivity matrix after learning**. The RNN connectivity matrix is initialized as a blank slate (all values are set to 0). After learning, the RNN connectivity matrix displays the appearance of clusters, whereby groups of 20 neurons are fully interconnected with each other and not connected with other neurons in the RNN (please refer to S2 Video for better visualization of clusters and their transitions as the sequence unfolds). Color bar represents weight values. **E. RNN i^{th} cluster learns to project to j^{th} Go node (see Fig 1, 'Detailed architecture', for the meaning of the indices)**. The top panel shows the randomly initialized weight values between the RNN excitatory units (before learning). The bottom panel shows how each cluster (represented by a subset of RNN neurons) is connected to a specific Go node after learning. Color bars represent weight values. **F**. Dynamics of G (left panel) and N (right panel) nodes after learning.

In simulation 3, we inject an additional input of +1 or -1 to the first Go node during the first 100 ms of the 1 second time window. Fig 3B shows how the sequence is shifted earlier in time for the positive input (left panel) and later in time for the negative input (right panel). Moreover, Fig 3C shows that as this additional input lasts longer, the distance (in time) between the first action of the shifted sequence and that of the original sequence increases linearly.

Simulation 4: Temporal rescaling. Musicians can learn a rhythm, i.e., a precisely timed action sequence, and instantly temporally rescale (compress or dilate) that rhythm without additional learning. In our model, flexible rescaling is achieved by sending a multiplicative input (ρ) to all Go nodes simultaneously (i.e. the ρ parameter multiplies the net input in Eq 5, see Methods); if $\rho > 1$ or $0 < \rho < 1$ the sequence is respectively compressed or dilated. Fig 3D shows temporal rescaling for ρ values of 1.2 (compression, left panel in Fig 3D) and 0.9 (dilation, right panel in Fig 3D). Importantly, temporally rescaling the sequence does not affect the temporal structure of action sequences. For 100 values of ρ, ranging from 0.9 to 1.2, we computed the relative ratio between a sequence of 3 actions. The ratio was computed by subtracting the time of action 1 from that of action 2 (subtraction 1), then the time of action 2 from that of action 3 (subtraction 2), and dividing subtraction 2 / subtraction 1. We performed this computation for the action triplets 1-2-3, 2-3-4, 3-4-5 and 4-5-6, and summed the ratios. Fig 3E shows that this sum of ratios stays constant (mean = 7.5, s.d. = 0.12), thereby indicating that temporal structure is maintained albeit rescaled. Note that rescaling also induces a tiny shift in the sequence. This is an emergent property of a global rescaling signal to all Go nodes of the network; this slight shift could be avoided by targeting all but the first Go node with the multiplicative term.

Simulation 5: Temporal compositionality. Musicians must also be capable of temporal compositionality; that is, applying a desired tempo to an action sequence that was learned in a different tempo (e.g., apply a bossa nova tempo to a rock song; see below). In simulation 5, we assume that desired tempos are learned and extracted from other sequences, which then can then be used as a dynamical multiplicative input signal to all Go nodes (Fig 3F right panel). Therefore, whereas temporal rescaling makes use of a constant ρ multiplicative input value to the Go nodes, temporal compositionality is achieved by a dynamic multiplicative input (the ρ multiplicative input value follows the blue trace in the right panel of Fig 3F). The result is to produce the learned sequence (described in Fig 3A) to the tempo described by the multiplicative signal. Fig 3F (left panel) shows how the time of each action in the sequence does not accord with the learned tempo (color coded vertical dashed lines), but is rather produced at the novel desired timing (vertical solid lines).

Simulation 6: Sustained motor activation. The ACDC model is also capable of producing sustained motor activation for any element within the sequence, for instance sustained notes in a musical scale. Our model can achieve sustained motor activation via two mechanisms. First, via a flexible mechanism similar to that of rhythm compositionality, a multiplicative signal ($\rho = 0.1$) is sent to the Go node during the period in which sustained motor
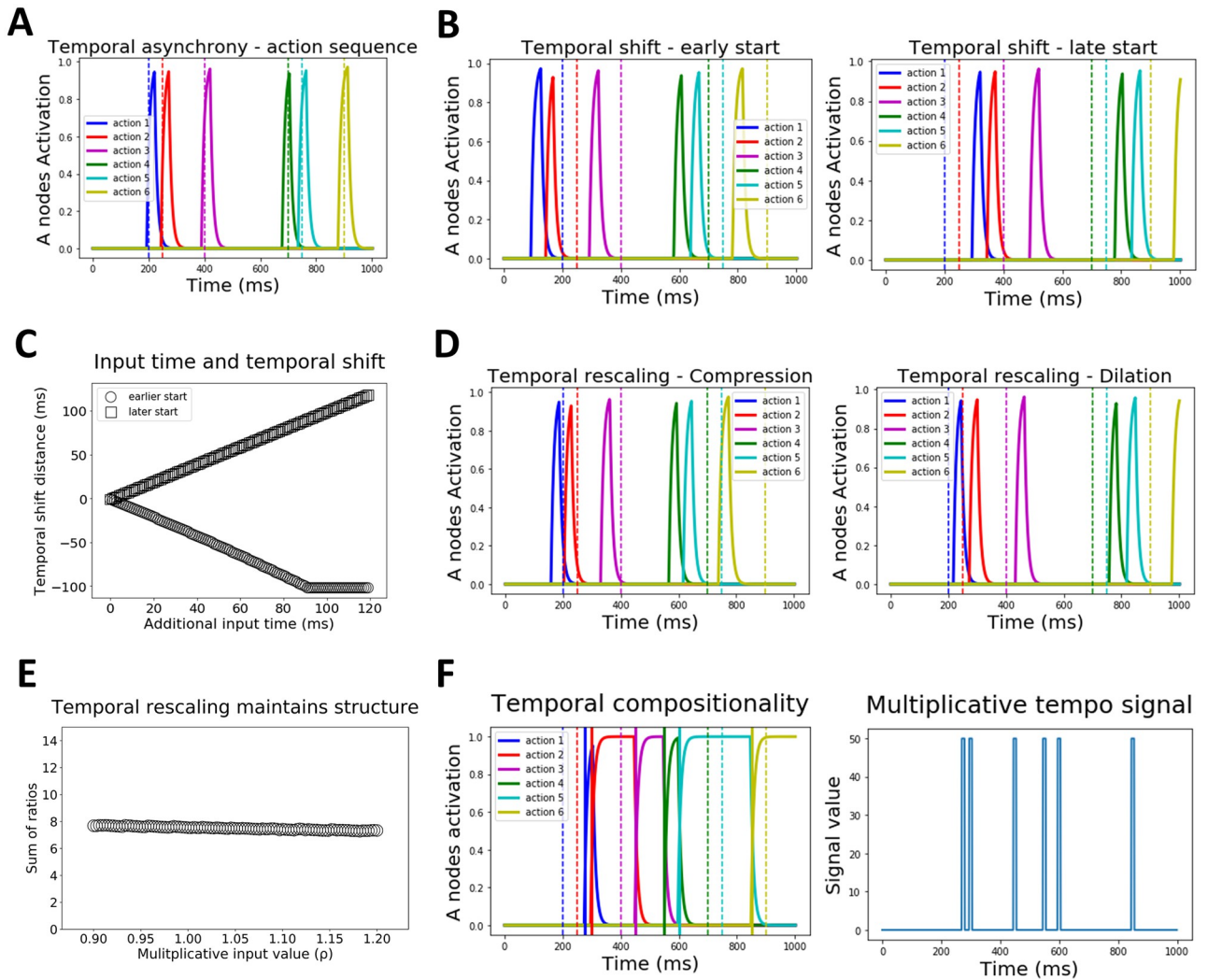
Fig 3. Temporal properties of the ACDC model. A. Simulation 2: Reproduction of action sequence with temporal asynchrony. Each action (i.e. A node activation, color coded) is produced at the precise desired time indicated by the vertical dashed line (also color coded), within a 1 second time window. Inter-action interval varies as the sequence unfolds. **B. Simulation 3: Temporal shifting**. A precisely timed action sequence can be started earlier (left panel) or later (right panel) by respectively injecting an additional positive or negative input to the first G node (i.e. associated to accumulating evidence in favor of the first action). Importantly, the temporal structure of the action sequence is not altered. **C. Simulation 3: Temporal shifting varies linearly with additional input time**. Applying longer input times leads to increasingly earlier or later shifts in sequence initiation times, depending on whether additional input is positive (circles) or negative (squares). **D. Simulation 4: Temporal rescaling**. Action sequences can be compressed (left panel) or dilated (right panel) by adding a multiplicative input to all G nodes simultaneously. **E. Simulation 4: Temporal rescaling preserves action sequence structure**. Importantly, when temporal rescaling is applied to the action sequence, the relative timing between each action (i.e. the structure) is preserved. Here, we plot the sum of ratios (y-axis, see main text) as a function of the multiplicative input ρ (x-axis). The sum of ratios value (black circles) stays constant as a function of ρ, indicating a preserved temporal structure even though the sequence is rescaled. **F. Simulation 5: Temporal compositionality**. The left panel shows how A nodes activity are activated on the tempo described by the multiplicative signal (left panel). Vertical dashed and solid lines on the left panel indicate the timing of each action for the previous and novel tempo respectively. As shown, the respective A nodes become active on the novel tempo.

activation is needed. Fig 4 left shows that applying such a signal between the second and third actions allows motor activation of the second Action node (red trace) to be sustained until the third action is executed (purple trace). Second, via a learning mechanism, the weight value between a specific Action-NoGo node pair can be decreased to induce sustained activation of the Action node (Fig 4, right). Note that we focus here on the mechanistic properties of the

Sustained motor activation
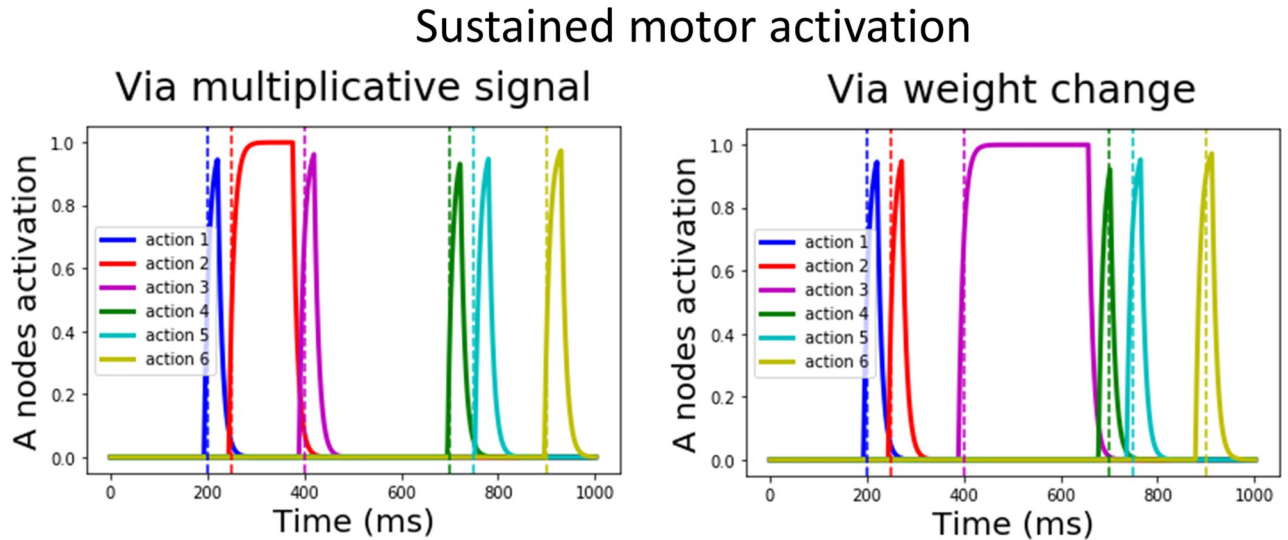
Fig 4. Simulation 6: Sustained motor activation. Both panels demonstrate that the ACDC model is able to output sustained motor activation as desired within a sequence. The left panel shows the results of applying a multiplicative signal ($\rho = 0.1$) to the second No Go node, inducing a sustained activation of the second action (red trace). The right panel shows a similar effect this time by decreasing the value of the Action-No Go connection of the third action, in turn inducing sustained activation of the third Action node (purple trace).

model, rather than proposing how the Action-NoGo weights may be learned to support sustained activation.

Simulation 7: The ACDC model in action and sound. Here, the ACDC model learns to produce the second guitar riff of ACDC's (the rock group) Thunderstruck song. This riff is composed of 16 actions hitting six different notes (B5, A5, G#5, F#5, E5, D#5) following an isosynchronous rock tempo (Fig 5A). By allowing the model to record each note corresponding to each sequential action (following Fig 5A), the ACDC model was able to musically reproduce the riff (S1 Audio file). Notably, S1 Video shows that the RNN dynamics represent sequential attractor states, encoding order and leading to the production of each action (and sound) in the sequence (for a slowed down demonstration of similar dynamics with a less complex action sequence see S2 Video below). Next, we leveraged temporal compositionality to allow the ACDC model to play the riff but now based on a bossa nova tempo without further training
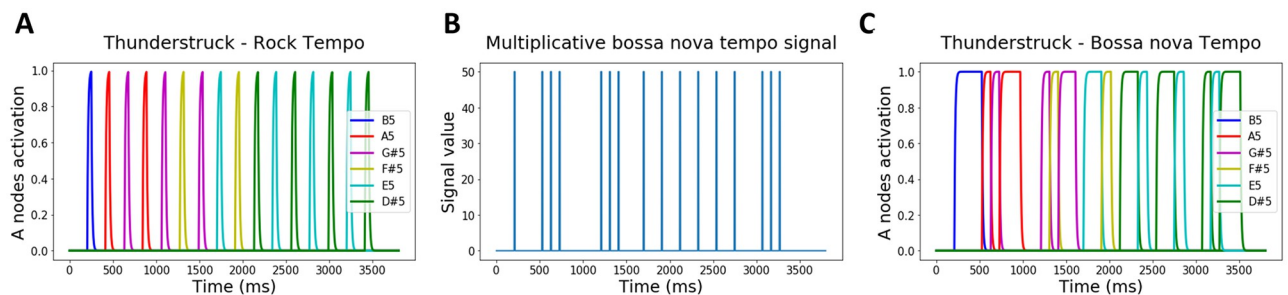
Fig 5. Simulation 7: the ACDC produces the Thunderstruck song. A. Second guitar riff from ACDC's (the group) Thunderstruck song. The riff is composed of 16 sequential actions creating a isosynchronous rock rhythm over a window of 3500 ms (given a 140 bpm tempo). Each action is associated to a color coded note). **B. Generic bossa nova tempo**. We imposed the model to replay the thunderstruck rock tempo song following a bossa nova rhythm whose tempo is described by the blue trace multiplicative signal. **C. Flexible generation of the Thunderstruck song following a bossa Nova tempo**. When the multiplicative input (Fig 3B) is given to the Go nodes of the BG, the ACDC model flexibly reproduces the Thunderstruck song but now following the bossa nova tempo.

([Fig 5B and 5C](#) and [S2 Audio file](#)). Further note that, altogether, this simulation encapsulates distinct temporal flexibility properties. First, flexibly reproducing the Thunderstruck song following a bossa nova tempo requires the ability to generate an action sequence with temporal asynchrony. Second, temporal rescaling is applied to parts of the song as the sequential execution of consecutive notes need to be sped up or slowed down. Third, the model displays its ability to produce sustained motor activation (see [S2 Audio file](#)).

### Behavioral and neurophysiological simulations

Simulation 8: Behavioral simulation. In the motor timing literature, a ubiquitous finding is *scalar variability*: when asked to produce an action after a specific time interval, the variability in action execution timing increases with the length of interval timing [63–66]. In simulation 8, our model learns to produce a single action at distinct interval timings (i.e. 200, 400, 600 and 800 ms). For each timing, the model produces 500 reaction times (RTs), from which we extract the standard deviation (SD), and reproduce this process for 100 simulations and two noise values (gaussian random noise with zero mean and SD of 0.01 or 0.05 is added to the model Eqs [3–5](#) and [7](#)). Reproducing empirical patterns, [Fig 6](#) shows that the SD of RTs increases as a function of interval timing, and thereby demonstrates that the ACDC model displays scalar variability (see also [67]). Furthermore, the SD value range also increases with noise values. This effect is explained in our model by having a fixed negative bias on the Action nodes in the motor layer. Such a feature reduces to having an accumulation-to-bound process for action execution. Hence, given a specific amount of noise, longer RTs are associated to wider RT distributions (i.e. larger SD, [68]). The underlying reason is that the effect of noise on evidence accumulation is amplified as time elapses.

Simulation 9: Neurophysiological simulations. Two other ubiquitous findings are persistent and sequential neural activity. First, several studies have observed persistent neuronal firing rates in temporal [69–71], parietal [72–75], premotor [76] and prefrontal [77–79] cortices whenever an agent has to hold in working memory task-relevant stimulus features (e.g., spatial location). Theoretical work suggests that persistent activation patterns emerge from recurrently connected networks that settle in one of multiple potential attractor states [80–82]. Second, as motivated in the introduction, sequential activity has also been observed in distinct sequential behaviors such as spatial navigation [8] and bird song [83–86].
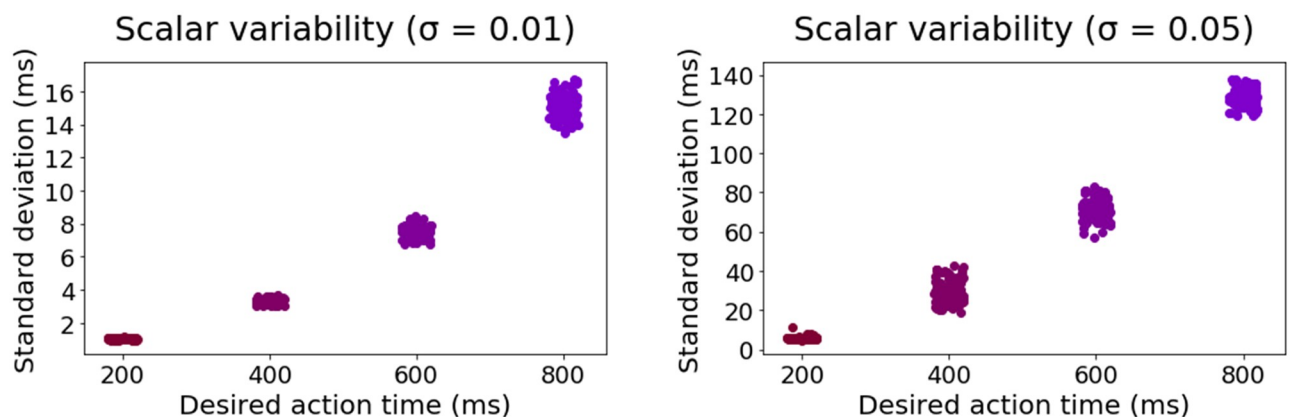
Fig 6. Simulation 8: The ACDC model displays scalar variability. Left (low noise value = 0.01) and right (high noise value = 0.05) panels show that the standard deviation of RTs increases as a function of the desired action time (i.e. interval timing). Moreover, higher noise values increase the range of standard deviation. Each dot is the result of 1 out of 100 simulations for each interval timing.
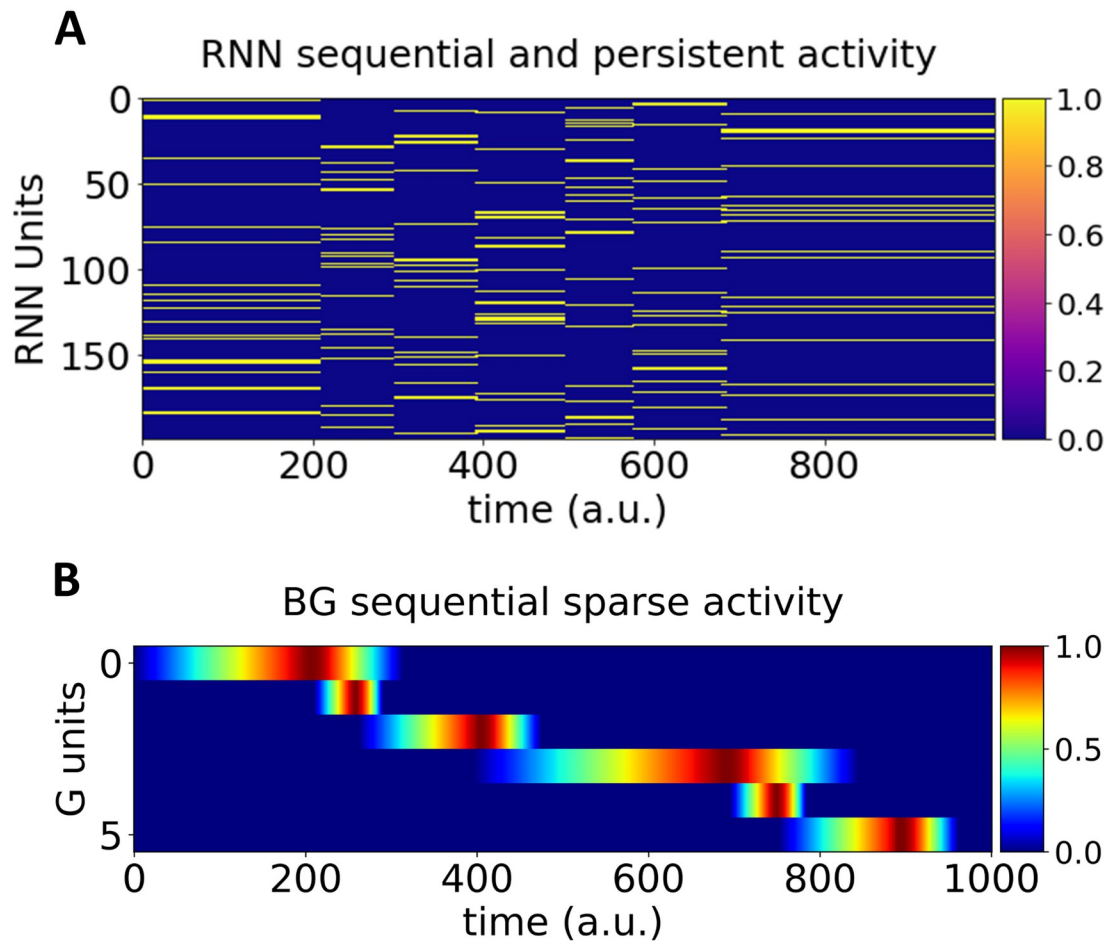
Fig 7. Simulation 9: A. Sequential and persistent activation of clustered neural populations within the RNN. The y-axis represents each RNN unit, the x-axis represents time. The first cluster is activated by the input layer, and maintains persistent activity until the first action is executed. At that moment, via excitatory projections from the Action nodes (Fig 1C) to the following (i+1th) cluster in the RNN (Fig 1B) gets activated, and thus displays persistent activation, and so forth via the cortico-basal ganglia loops (light blue arrows in Fig 1). Color bar represents firing rate. **B. Sequential and sparse activation in the BG**. The y-axis represents the G unit activity over time (x-axis). Each G unit responds in a sequential and transient manner, as has been shown in neurophysiological single-cell recordings of the BG (e.g., [4]). Color bar represents normalized firing rate.

https://doi.org/10.1371/journal.pcbi.1009854.g007

Interestingly, recent work suggests that sequential switches in attractor states (and hence persistent neural activity) occur during behavioral switches in action sequences [87]. Therefore, persistent and sequential activity may emerge from the same mechanism. In our model, the RNN activation dynamics display such switches from one attractor to another as the action sequence unfolds. Each attractor state is associated to the persistent activity of neurons forming a cluster in the PMC (RNN). When the action associated to that attractor state (i.e. the jth action associated to the ith order) is executed, this triggers a switch in attractor state in the RNN (via cortico-cortical projections from M1 to PMC), as empirically observed [87]. In simulation 1, the ACDC learns to produce an arbitrary sequence of 6 actions, each with their own desired execution time within a window of 1 sec (i.e. at 200, 250, 400, 700, 750, 900 ms). Fig 7A shows the RNN dynamics after learning. Each cluster of activation displays persistent neural activity until the action is executed, which triggers the following cluster of persistent neural activity. Hence, activity in the RNN is both persistent and sequential in nature.

To gain better visual intuition on the RNN dynamics, we performed dimension reduction on the row space of the unit (i.e., neuron) by time matrix displayed in Fig 7A. We then dynamically plotted the first 3 principal components (PCs) as a function of time. S2 Video shows that each cluster of persistent neural activity acts as an attractor state (within the highly dimensional space of the RNN), and the dynamics in the RNN switch from one attractor to the other when an action is executed, again displaying both persistent and sequential neural dynamics.

The qualitative pattern of the RNN sequential and persistent dynamics (Fig 7A) is different than observed in rodent [2,8,11] or monkey [1] neurophysiological recordings, which reveal sequential sparse activation (individual neurons display quick and transient activation as behavior unfolds). Notably however, the Go nodes in the BG module of our model display qualitatively similar sequential and sparse activation patterns as that seen empirically in the BG (Fig 7B; see Figs 2A, 3B, 1E and 8C, respectively of [2,3,5,88]).
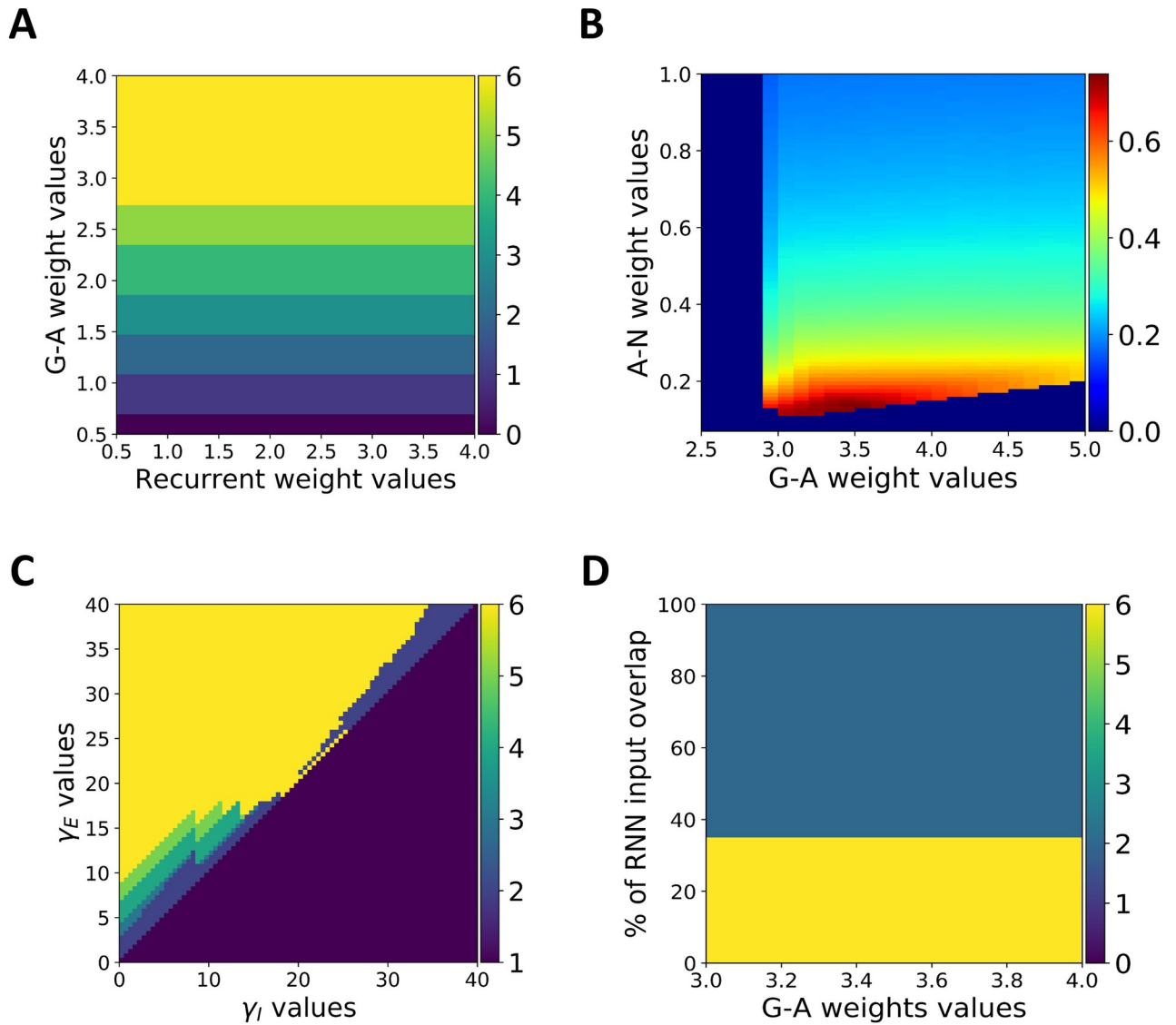
Model regimes and robustness

A useful model should be robust to variations in its key parameters and/or should exhibit regimes in which the model exhibits qualitatively different features [23,89–91]. To examine the sensitivity of our main results to parameter variations, we start by exploring the model regimes in a network with fixed connectivity. We focus on the recurrent weights within the excitatory RNN units, as well as the weights from the G to A units, which we consider similar to recurrent and feed-forward weights in traditional associative chain models [24]. We explore the necessary weight combinations for the model to execute an entire sequence of 6 actions, within a temporal window of 1 sec (as described in the learning subsection). Fig 8A shows that weight values from G to A units subtend the number of actions that can be executed within the 1 sec temporal window: higher G to A weights progressively lead to sequences with more actions. Notably, recurrent weight values do not influence the model's ability to produce the sequence (no impact of variation along the x-dimension). In contrast to previous models [23], recurrent and G to A weights (i.e., feed-forward weights) do not control the model regimes, i.e., whether the model produces damping, sequential or persistent activity. This difference emerges from divergent architecture between our and previous models. Persistent activation of each action during the entire sequence cannot emerge in our model. Indeed, action execution is automatically associated to a switch in attractor state within the RNN. This hinders the previous state to continuously activate its G unit and hence cannot provide evidence for the previous A unit. Therefore, if any, persistent activation of an action can be maintained only up until the next action is executed.

As shown in simulation 6, one way of controlling persistent activity is through the weight values of the A to N unit projections: activation of a particular action shuts down that action via feedback projections to N units, thereby reducing persistent activation. To quantify persistent activation across an action sequence, we derive a "sustainability" measure by computing the area under the curve (AUC) for each A unit activity that falls above the value of 0.5, summing the individual AUCs, and normalizing by the total sequence time (i.e. 1000 ms). Fig 8B shows a heatmap of this measure as a function of A-N and G-A weight values. Sustainability increases with lower values of A-N weight values (note that we plot this measure only for parameter combinations that produced a full six action sequence; otherwise we set this value to 0). Interestingly, the entire range of sustainability is covered by the A-N weight values over a broad range of G-A weight values, from transient pulse (blue values) to sustained actions (red values).

As noted above, a prominent feature of our model is that action execution triggers switches in RNN attractor states through parallel projections to its excitatory and inhibitory units. For

Fig 8. Model regimes and robustness. A. Number of actions within a sequence as a function of feedforward and recurrent weights. The y- and x-axis represent G-A and recurrent weight values, respectively. As depicted, the G-A weight values control whether the model can produce six actions within a 1 sec temporal window, irrespective of the recurrent weight values. The color bar codes the number of actions that are produced within the sequence; yellow for a full sequence (six actions) and dark blue for no actions. **B. Action sustainability**. The heatmap reflects the sustainability measure magnitude (warm colors coding for higher values) as a function of A-N and G-A weight values. As depicted, action sustainability increases with decreasing values of A-N weights, over a large range of G-A weight values. **C. Model regime as a function of the γ_E and γ_I parameters**. The y- and x-axis represent γ_E and γ_I parameter values, respectively. As depicted, the model can reproduce fully and precisely a learnt action sequence within a broad range of parameter combination respecting the $\gamma_E > \gamma_I$ inequality. Color bar is identical to A. **D. RNN input overlap**. The model can produce a full six action sequence up to 35% of input overlap to the RNN between contiguous actions; i.e. activating 35% of the previous and subsequent cluster in the RNN. Stronger overlap leads to a break in the sequence after 2 actions. Color bar is identical to A and C.

https://doi.org/10.1371/journal.pcbi.1009854.g008

the switch to take place, the gain must be stronger on projections to RNN excitatory vs inhibitory units (i.e. $\gamma_E > \gamma_I$), allowing the previous cluster to shutdown while also activating the next cluster in line. To explore the robustness of this property, we used the learned sequence in simulation 1, and parametrically modulated both gain values (γ_E and γ_I values are represented on the y and x dimension, respectively). As depicted in Fig 8C, a vast range of gain

combinations allows the model to fully and precisely reproduce the learnt sequence (yellow space), as long as $\gamma_E > \gamma_I$. Note that some combinations allow for a partial reproduction of the sequence.

Finally, one simplifying assumption in the model is that feedback inputs to the RNN are orthogonal and do not overlap (i.e., projections from distinct A units never activate the same RNN excitatory units). In Fig 8D we relaxed this assumption and progressively increased the percentage of overlap between contiguous A unit inputs to the RNN. We observed that the model can continue to produce a complete six action sequence as long as the overlap between inputs of contiguous actions does not surpass 35%, at which point the model fails to reproduce a full sequence and only manages to produce the two first actions as two quick pulses. Further, note that the effect of input overlap is independent of the G-A weight values (taken from a range that produces a full sequence, see yellow area in Fig 8A).

Discussion

The ACDC model combines the strength of associative chains (e.g., [23]) and cluster-dependent (e.g., [27]) models, while also formalizing how the BG contribute to recurrent cortical dynamics in sequential behaviors. Our model factorizes action order, identity, and time, which are represented in distinct loci of the cortico-basal ganglia neural network. Crucially, factorizing these features provides the network with the ability to independently manipulate the building blocks of precisely timed action sequences, thereby increasing the computational power of our model. This increased power is illustrated through several interesting emergent properties. First, we demonstrated that the ACDC model can learn and reproduce precise spatiotemporal action sequences with temporal synchrony or asynchrony. Second, our model displays several flexibility properties: temporal shifting, rescaling and compositionality, and sustained motor activation; culminating in our model's ability to reproduce the Thunderstruck song and change it to a bossa nova tempo. Third, the model can account for behavioral and neurophysiological empirical observations. Finally, we showed that the main model properties are robust across a range of parameter values.

Encoding order as attractor state switches in the RNN

Recent work suggests that dynamic representations can be understood as switches in activity of neural networks [92], whereby action sequences emerge from neural attractor states unfolding over time. Indeed, Recanatesi et al. [46] showed that sequential behavior was subtended by the sequential unfolding of attractor states in rodent secondary motor cortex. Furthermore, these authors modeled variability in action timing by adding correlated noise to the dynamics of a RNN, leading to dynamics that jump from one attractor state to another at random times (hence explaining the variability in action timing). In our model, the switch between cortical attractor states is not random but is controlled by dynamic BG modulation of excitatory thalamic projections to the RNN that transiently modify the ratio of excitatory to inhibitory inputs. Within our conceptualization, we suggest that persistent activity within a cluster indicates the latent state that the system is in [90], which in this case reflects the ordinal position in the sequence. Moreover, in contrast to previous models of frontal cortical BG interactions [90], the cortical clusters themselves were not assumed to be anatomically hard-wired but emerged within the RNN via learning.

Alternative models have proposed different mechanisms for encoding ordinal position. Some models possess a temporal context layer whose state is modified dynamically as time passes [93–96]. Other models assume that the network input (used to learn the sequence) is itself sequential in nature [27,28], and learning the spatiotemporal signal depends on the

sequential nature of the input. Our model is free of this assumption; the network input is a single pulse of activation, but can nevertheless reproduce a precisely timed spatiotemporal signal. This ability emerges from the feedback loop from thalamic Action nodes to the cortical RNN, triggering transitions to a subsequent cortical attractor. One can therefore consider motor output as part of the teaching input signal to the RNN; because motor activation unfolds sequentially in our model, the sequential nature of the teaching signal emerges from our network architecture.

Interestingly, the idea that the motor cortex (presumably via motor thalamus neurons) acts as a teaching signal to other brain areas has received strong support from rodent lesion studies. For instance, rats are unable to learn a precisely-timed lever press when their M1 cortex is lesioned [97], and transiently inactivated or disturbed via optogenetic manipulation [98]. More generally, the notion that motor output can influence cognitive representations and transitions is consistent with the emerging literature on how cognitive functions scaffold on top of motor functions in cortico-basal ganglia circuits [99,100].

Motor sequence flexibility as inputs to the basal ganglia

Humans can adapt their motor output almost instantaneously given external or internal stimuli. For instance, musicians can modify the tempo of a song upon signaling of the conductor. Such flexibility necessarily needs to stem from fast reconfiguration of neural dynamics, rather than emerge from changes in networks weights [12]. Murray and Escola [28] proposed a model of interconnected medium spiny neurons in the striatum that can apply such dynamic reconfiguration. In particular, their model could perform temporal rescaling of sparse sequential activity. Yet, flexibility in this model is constrained to isosynchronous sequences (see also [67,101]). However, a recent model [22] making use of eligibility traces [102–106], manages to learn precise asynchronous spatiotemporal sequence learning. Still, it is unclear how such a model can rescale asynchronous sequences, and neither of these models is capable of exhibiting temporal compositionality. Nevertheless, ring-like models with synaptic depression [e.g., 28] could potentially account for these properties. Indeed, temporal rescaling in these models is often implemented as changes in the background current (higher current levels lead to faster rescaling). Therefore, to produce asynchronous sequences, one could imagine a dynamical background current which is null when no action has to be provided, resulting in a silent network, and turned on when the sequence has to be resumed. However, given a silent noisy network, reactivation of a global current will induce activation of a random cluster in the sequence and not necessarily the next cluster in line, thereby not displaying robustness in sequence production. Indeed, for such a system to be viable, the network would need an additional piece of information, which is a memory of which cluster (or unit) was last active. The reason is that, in ring-like models the background current is global (i.e., sent to all units). Thus, if the network goes silent, reactivation of the network would not ensure that the appropriate cluster (i.e. the one next in line) becomes active; i.e. noise would randomly activate a cluster and the sequence would restart from that cluster (see implementation of [28]). Consequently, some extra input signal (representing a memory of which cluster was last active) should give an advantage to the next cluster in line, in order to ensure that sequence order is maintained. This extra input signal could take the form of recurrent connections within the cluster, controlling the activation decay of that cluster such that it would still be active after long periods of time (as opposed to all other clusters that would be fully silent; see [22]). Note however that recurrent connections would put a break on the synaptic depression, as these two parameters trade off; therefore, formal implementation of this proposal would be needed to understand the model regimes.

Crucially, the ACDC model can perform temporal rescaling for both iso-synchronous and asynchronous sequencing, and it can also flexibly switch the tempo altogether through a multiplicative signal to the BG. Our model proposes a more robust approach to sequence flexibility. The sequence in our model is chained by the execution of the previous action, thus our model does not require an explicit representation of the most recent action. Indeed, what controls the sequence is a series of local signals (i.e., selective feedback from the motor thalamus to the RNN [49]) rather than a global signal, and what controls the timing is a global input to the BG. This feature is key for our model to account for the distinct types of flexibility described in the results.

It is important to clarify that flexibility in our model is implemented at the action selection level rather than action execution or implementation. Indeed, although we use musical examples to motivate our work, our model focuses on the timing/selection of actions rather than their execution. Even simple finger movements implemented to play the piano require muscle commands represented as highly dimensional and continuous signals [40,107,108].

The temporal properties of our model discussed above emerge from additional inputs to the BG. What is the nature of this input? One possibility could be dopaminergic. Indeed, midbrain dopaminergic nuclei massively broadcast to the striatum [109], and several studies have implicated dopamine in controlling movement vigor [110–118]. Dopamine has also been extensively implicated in impulsive (i.e. pathologically speeded) behavior [119–125]. Furthermore, administration of amphetamine and haloperidol to human participants, respectively increasing and decreasing tonic dopamine levels, has been associated to faster and slower response times during a simple reaction time task [126].

If dopamine can flexibly modulate (speed up or slow down) action execution timing, the question remains upon which psychological process this neuromodulatory effect takes place. Within the accumulation-to-bound framework [68,127], this effect could potentially alter two distinct processes. First, dopamine could play a role on the speed (or rate) of evidence accumulation. In line with this hypothesis, several studies have highlighted a clear effect of dopamine on the drift rate of evidence accumulation in perceptual [128,129] or reward-based [130] decision-making tasks. Our model implements this possibility. Indeed, inputs to Go nodes modify (increase or decrease) the drift rate of evidence accumulation. Yet, the speed at which an action is produced also depends on the response threshold, with lower thresholds increasing speed at the expense of accuracy [131]. Therefore, a second alternative is that dopamine or other BG modulations may modify the threshold of action execution [132,133]. Interestingly, Parkinson's disease patients on subthalamic deep brain stimulation tend to behave impulsively [125], due to modulation of the decision threshold [134–136]. Naturally, both hypotheses are not mutually exclusive; further research should investigate the effects of dopaminergic and subthalamic modulations regarding motor sequence flexibility.

Limitations and future directions

As previously noted, some of the implementational and biological details of our model remain to be worked out. First, we simplified the BG gating circuitry, to focus on the G and N populations, summarizing their effects on downstream thalamus but omitting the disinhibitory circuitry involving the substantia nigra. Many previous models, including our own, have simulated the more complete direct and indirect pathways but we did not feel this detail was necessary for the present purposes. Second, reinforcement learning of action timing is conceptually thought to take the form of a three-factor hebbian learning rule [43,103,137–139], where neurons subtending a rewarding behavior (and hence forming a specific cortical activity patterns) increase their connectivity to D1-receptor containing striatal "Go" populations via

dopaminergic activity bursts stemming from midbrain nuclei [90,140–143]. While we do not challenge this mechanism, we focused learning in our model to one synapse downstream, from Go nodes to thalamic motor neurons. Indeed, recent evidence suggest that error-driven learning can be achieved via manipulation of BG outputs to thalamus [144]. Third, although much evidence indicates that the BG learn via reinforcement learning [145–147] (i.e., depending on whether rewards are better or worse than expected), we incorporated a signed error in our learning mechanism which is more powerful for timing signals. The supervised delta rule is only used at the output of our network to optimize readout weights. Thus, we do not view this learning rule as implausible in our model given that we focus on sequence learning situations in which signed error feedback is provided (see [61]), such as when a tutor teaches you how to play the drums and holds the tempo, or in bird-song learning [84] (where a tutor is available to provide signed error). There are several biologically plausible implementations of the delta rule when such error signals are available (e.g., [148,149]). Our learning rule in the BG-thalamus thus summarizes the contributions of these systems in conjunction. Nonetheless, future work should investigate how and whether complex precise sequences may emerge solely based on reinforcement learning. In principle, the ACDC model could learn sequences solely based on RL. However, learning would be much more tedious [43]. In contrast to ACDC, other models have directly trained the entire time-series of individual RNN units to match empirical data [38,150], using continuous learning signals (also see [34,36,37,39]). Hence, the neurophysiological simulations in ACDC (i.e., sequential sparse activation and attractor states) emerge only from the proposed theoretical architecture in the context of behavioral experience. Fourth, we implemented a single inhibitory neuron in the RNN-PMC module. Our focus was on the functional role of inhibitory neurons on the transition of attractor states [28], and certainly this single unit could be replaced by a larger population. Future versions of the model should include a broader pool of inhibitory neurons in the RNN as they have been shown to exhibit mixed selectivity to multiple aspects of a task [151]. Fifth, one assumption in our model is orthogonality of all projections to the RNN. Although we showed that this assumption can be relaxed to a certain extent, this feature ensures there is no ordinal interference during action sequence execution. In the brain, this orthogonality may be implemented via mixed selectivity of excitatory frontal neurons that ensure downstream readouts without interference [152,153]. Interestingly, Márton et al. [154] recently developed a RNN model of cortico-striatal interactions optimized to learn oculomotor sequences. Similar sequences were performed by awake monkeys while activity was recorded in their dorsolateral prefrontal and striatal areas. Learning to implement the correct actions for each sequence pulled apart the representational structure of action sequences in activity space both in the model and neuronal recordings. Whereas ordinal representations in our network were hardwired as orthogonal vectors in the RNN (in order to avoid interference), the work of Márton et al. suggests this may emerge naturally through learning.

Our model simulates action sequences such as those needed to play the guitar or the piano. Within this context, each action is represented as a discrete entity. However, many daily life action sequences are subtended by more continuous actions, as for instance when playing violin with a bow. The ACDC could be expanded by having more continuous representations of action plans and execution in our BG-thalamus module. Based on dynamic field theory, one potential approach would be to represent actions as dynamic neural fields [155–157], which have been shown to successfully model more continuous reaching actions [158]. Moreover, these continuous action representations in the BG may require additional inputs from the cerebellum for movement coordination [159] or sequence prediction for motor control [160]. Moreover, our model (as others [22,28,67]) was specifically engineered to account for spatio-temporal sequences and how these may be flexibly manipulated. This in contrast to other

instantiations [35,161] of RNNs (i.e. reservoir computing) that find natural solutions to diverse tasks involving distinct psychological processes (e.g., memory, time estimation, decision-making).

Finally, recent research focused on how humans extract abstract knowledge, and generalize this knowledge to other situations [90,162–164]. Indeed, abstracting the action sequence structure of the Thunderstruck song may be useful for future learning. Transferring the abstract structure of the Thunderstruck song when learning a novel song that shares a similar structure should improve learning [165].

Methods

Below we provide a full description of the ACDC model; parameter values for all simulations are reported in Table 1, and code is available from https://github.com/CristianBucCalderon/ACDC.

The associative cluster-dependent chain (ACDC) model for flexible motor timing

Our ACDC model contains four main modules (Fig 1): an input layer (Fig 1A), an RNN (representing premotor cortex; Fig 1B) and a BG-thalamus unit (Fig 1C).

The input layer (Fig 1A) consists of a vector of neurons, of which a subset is activated, representing sensory or other context that would signal the identity of the sequence to be produced or learned.

Crucially, the dynamics within the ACDC model evolve as a sequential unfolding of RNN-BG-thalamus-RNN (i.e., cortico-basal ganglia) loops, depicted by the light blue arrows in Fig 1. The sequence starts with the activation of a cluster (i.e., densely interconnected) of excitatory RNN neurons (Fig 1B). Each cluster will come to encode the ith element in the action sequence. As opposed to single unit, clustered neurons provide a biologically plausible mechanism for supporting persistent activation within the cluster given a phasic input (i.e., an attractor, [81,166]). In prefrontal cortical–BG models, such clusters are referred to as "stripes" based on their anatomical existence, and are independently gated by BG [167]. Once a cluster is activated, the RNN temporarily settles on an attractor state indicating the ordinal position (order or rank) in the sequence, analogous to how distinct PFC stripes code for ordinal

Table 1. Parameter values for all simulations.

Parameters	Values
α_1 (RNN) / α_1 (RNN-Go)	0.01 / 0.00002
α_2 (RNN) / α_2 (RNN-Go)	0.1 / 0.4
W_{max} (RNN) / W_{max} (RNN-Go)	1 / 0.05
b	0.5
τ_w	2
η	0.4
ϕ	0.01
J^{IE}	0.1
τ_a, τ_n	10
γ_E / γ_I	21.4 / 21
$\tau_{rnn}, x^{in}, J^{EI}, J^{EA}, J^{IA}, J^{GN}, J^{NA}$	1
λ_{rnn} / λ_a	10 / 10000
τ_g	1000

https://doi.org/10.1371/journal.pcbi.1009854.t001

positions in phonological loop tasks [167]. However, in ACDC such clusters emerge naturally via learning rather than hard-coded anatomical entities. Moreover, attractor states are maintained via a specific ratio of excitatory to inhibitory inputs: each excitatory neuron projects to a common single inhibitory neuron (orange circle in Fig 1B) which reciprocally inhibits all excitatory RNN neurons. As long as the ratio of excitatory to inhibitory inputs is not perturbed by another input (see below), activation in the cluster will persist and the RNN will continue representing the ith order in the sequence.

In turn, each excitatory RNN cluster projects to its corresponding "Go" unit in the BG (blue arrow 1 from i^{th} cluster in Fig 1B to G node in Fig 1C), and each Go cell accumulates evidence for the jth action associated to the ith order (see [134,168] for related computational models of evidence accumulation in these units, and [169] for empirical data). Striatal Go cells, via the basal ganglia direct pathway machinery [170,171], facilitate response execution by projecting towards the corresponding motor thalamus neurons, from here on termed Action nodes for simplicity (blue arrow 2 from Go to Action nodes in Fig 1C). The BG component summarizes the contributions of more detailed BG circuitry [140,167,172]. In these models striatal neurons accumulate evidence, which via the direct and indirect pathways leads to categorically discrete signals in BG output nuclei, and to disinhibition of the thalamus (e.g., [132,168]). These patterns are also observed empirically in terms of striatum accumulation signals and discrete downstream responses in BG output nuclei once a threshold of accumulation is reached [169]. Here, we lumped together the double inhibition from striatum to Globus Pallidus (GP) and from GP to the thalamus into a single excitatory projection to keep the model simple and tractable. Interestingly, optogenetic stimulation of the GP has been shown to increase the firing rate of motor thalamus neurons [173].

Action nodes possess a negative bias, which acts as a decision threshold, i.e., the net input needs to exceed this bias in order for action to be executed. This feature again summarizes the computational role of the output of the BG, which serves to inhibit action execution until sufficient evidence reaches the threshold for action gating ([132,134]; see also [174]). Therefore, the weight values between Go and Action nodes control the speed of action execution: the BG encode the rhythm. Action execution can be expressed either as a transient or persistent response (see simulations; [23]).

In turn, Action nodes project excitatory connections to three distinct parts of the network simultaneously. First, Action nodes project to the cluster of excitatory neurons in the RNN representing the $i+1^{th}$ order in the sequences (blue arrow 3a in Fig 1). Second, Action nodes project to the inhibitory shared neuron (blue arrow 3b to orange node in Fig 1), that in turn globally inhibits all the clusters in the RNN. In this manner, thalamic Action nodes can update the cortical representation by separately projecting to both inhibitory and excitatory neurons [52,175], enabling the RNN to transition from the current state to the next. That is, the activation of action nodes perturbs the ratio of excitatory to inhibitory RNN inputs in a way that allows the ith cluster to shut down and the $i+1^{th}$ cluster to be expressed. Third, Action nodes project excitatory connections back to their corresponding No Go cells (blue arrow 3c from j^{th} Action node in the thalamus to j^{th} No Go node in the BG, see Fig 1C). In turn, No Go cells strongly inhibit their corresponding Go cells [132,176,177], thereby shutting down evidence in favor of the jth action, and hence stopping the execution of the jth action. This loop is then reproduced with the $i+1^{th}$ RNN cluster and $j+1^{th}$ G-A-N triplet in the BG-thalamus unit, and so forth until the action sequence is performed in its entirety.

Several features of the model should be highlighted. First, each cluster activation within the RNN acts as an attractor state representing the ith element in the sequence. Interestingly, cells in the monkey PMC code for the position in sequence, regardless of the actual movement produced during that position [178–184]. We therefore assume that the neurons forming each

cluster represent rank-order-selective neurons whose activation unfolds sequentially: the RNN encodes order information.

Second, the speed at which each action is executed is driven by how quickly the evidence in the Go nodes of the BG can cross the decision threshold in the Action nodes: the BG encode time information. Indeed, several studies suggest that temporal processing is subtended by the BG in the (non)human primates and rodent brain [1,2,185–189]. Note that there are multiple routes by which timing can be altered within Go nodes in our model: (i) the learned weight value between Go and Action nodes; (ii) a bias input to Go nodes (in addition to that coming from the RNN cluster); and (iii) a multiplicative gain on Go unit activity (see model simulations). As shown in the results section, these separate routes are important for providing timing and rhythm flexibility.

Third, as in many cortico-BG models (e.g., [134,190]), and motivated by anatomical data [191] our model is characterized by topographical organization of actions across the BG circuit and its outputs (i.e., indexed in our model by the subscript j associated in the G-A-N triplet projections). Recent evidence further confirms topographical action representations in BG-thalamocortical loops [192–194], whereby causal activation of specific subregions is related to specific output behaviors [195], and is also supported by human neuroimaging [196] and monkey/rodent neurophysiology studies [1,197–202]. However, in contrast to previous models in which BG gating affords action selection of the corresponding cortical action, in the ACDC model BG gating triggers a cortical dynamical state that initiates the evolution of the *subsequent* item in the sequence.

Fourth, we clarify how the ACDC model combines properties of associative chain and cluster-based models. While the ACDC model does initiate a *chain* via sequential propagation across cortico-BG loops, the timing of such transitions is controlled by learning the weights within the BG-thalamus unit, and moreover, what is learned are transitions between clusters of excitatory RNN neurons representing order in the sequence [27]. Hence, the ACDC model makes use of two distinct conceptualizations of sequence learning, to achieve greater computational flexibility (as demonstrated in the result section).

Learning in the ACDC model: Hebbian learning for order and Delta rule for time

Learning in the ACDC model takes place in three distinct loci of the network, comprising Hebbian learning for sequence transitions and error-driven learning for precise timing.

First, as previously mentioned, order is coded via persistent activation within clusters of the RNN. However, in contrast to pure associative chain models, the ACDC does not assume any feedforward hard-wired structure, but rather learns it. Selective time-dependent inputs to the RNN (i.e., from the input layer and thalamic Action nodes) activate a subset of neurons within the RNN, which get clustered together through dynamic synaptic weights:

$$\frac{dW_{ij}}{dt} = -\alpha_1((1 - x_i)\bar{x}_j) + \alpha_2(x_i \bar{x}_j(W_{\max} - W_{ij})) \tag{1}$$

where $\bar{x}_j$ is presynaptic activity low-pass filtered over a time scale τ_w; x_i is postsynaptic activity; α_1 and α_2 are learning rate parameters. When $\bar{x}_j$ and x_i are both simultaneously > 0, W_{ij} goes to *Wmax*; otherwise W_{ij} decreases (note that we clamp W_{ij} such that $W_{ij} \geq 0$). Note that $\bar{x}_j(t)$ will be non-zero if unit j is active within the time window from $t - \tau_w \to t$. Note that low-pass filtering is not strictly necessary in our version of the model and was implemented to maintain consistency with previous work in the domain [28]. However, its value should be < 2 otherwise all RNN units will tend to be connected and action sequence learning will fail.

Second, Eq 1 is also used to learn connections between the RNN and the Go nodes of the BG module; here, pre- and postsynaptic activity refer respectively to RNN excitatory unit activity and Go nodes activity (weight values between RNN units and Go nodes are randomly initialized from a Gaussian distribution with mean = 0.5 / N and s.d. = 0.1 / N, where N is the number of RNN excitatory units).

Third, action specific execution time is coded in the weights connecting Go and Action nodes. Here, we describe time learning as a delta rule, whereby an agent receives a supervisory signal explicitly indicating whether a specific action has been produced before (positively signed signal to increase weights) or after (negatively signed signal to decrease weights) the appropriate time, and is described in Eq 2:

$$\Delta W = \eta(t_{observed} - t_{desired}) \tag{2}$$

where the change in weight (ΔW) between the jth Go and Action nodes is driven by the learning rate η, and the error computed as the difference between the observed and desired response time (t) for each action. Weight values between Go and Action nodes are randomly initialized and drawn from a random Gaussian distribution (mean = 2, s.d. = 0.2). Learning of precisely timed sequences is shaped sequentially: the model first learns to produce the first action at the appropriate time (i.e. until the error $< \varphi$ and φ is a low value, see Table 1), then the second, and so forth.

Mathematical description of the model dynamics

The input layer reflects a vector of $N = 200$ neurons of which a subset (20) is activated and each neuron excites only one neuron in the RNN.

The dynamics within the ACDC model represent the sequential unfolding of RNN-BG-thalamus-RNN (i.e., cortico-basal ganglia) loops, depicted by the light blue arrows in Fig 1. The loop starts with the activation of a cluster of excitatory RNN neurons, and the dynamics of the RNN excitatory neurons are governed by Eq 3:

$$\tau_{rnn}\frac{dx_i}{dt} = -x_i + \Theta\left(\sum_{j=1}^{N}W_{ij}x_j - J^{EI}x_I + J^{EA}(x_A\gamma_E) + x_i^{in}\right) \tag{3}$$

where x_i and x_j represent post- and pre-synaptic RNN unit activity (purple nodes in Fig 1B) and W_{ij} is the recurrent weight matrix. J^{EI} and J^{EA} represent respectively the weights from the shared inhibitory neuron (orange node in Fig 1B) and from the motor thalamus neurons (from here on termed Action nodes for simplicity) to the excitatory RNN units. x_I, x_A and x^{in} represent respectively the activity of the shared inhibitory neuron, Action nodes (see below), and the input to the excitatory RNN units. γ_E is the gain on Action nodes activation projected to the excitatory RNN neurons (see below for the functional property of this parameter). Θ, the non-linear transformation function, is governed by $\Theta(x) = (2 / (1 + e^{-\lambda x}))- 1$ (where λ is the gain parameter and with additional non-linearity at zero, i.e. $\Theta(x) = 0$ if $\Theta(x) < 0$); and τ_{rnn} is the encoding constant. Note that input projections and all Action nodes to RNN projections are orthogonal (i.e. some RNN excitatory neurons receive inputs from the input layer, whereas others receive from inputs from Action nodes; each projection excites 20 RNN units). The shared inhibitory x_I activation is described by Eq 4:

$$\tau_{rnn}\frac{dx_I}{dt} = -x_I + J^{IE}x_i + J^{IA}(x_A\gamma_I) \tag{4}$$

where J^{IE}, J^{IA} and γ_I respectively represent the weights from the excitatory RNN neurons to

their shared inhibitory neuron, the weights from the Action nodes to the shared inhibitory neuron, and the gain on Action nodes activation for the projections towards the inhibitory neuron in the RNN.

In turn, each excitatory RNN cluster projects to its corresponding "Go" cell in the BG (blue arrow 1 from Fig 1B to Go node in Fig 1C), and each Go cell accumulates evidence for the jth action associated to the ith order, following Eq 5:

$$\tau_g \frac{dg_j}{dt} = -g_j + \sum_i^N W_{ij} x_i - J^{GN} n_j \tag{5}$$

where g_j is the activation of the jth Go units, W_{ij} is the weight matrix representing connectivity between RNN and Go units, x_i is the acitivity of the RNN excitatory units, J^{GN} is the inhibitory weight between the jth No Go and Go nodes, n_j is the activation of the jth No Go node, and τ_g is the encoding constant (with $\tau_g >>> 0$, thereby simulating evidence accumulation-like dynamics). Non-linearity at zero is also applied to Go-units.

Striatal Go cells facilitate response execution by projecting towards the corresponding Action nodes (blue arrow 2 from the Go to Action nodes in Fig 1C), whose dynamics are governed by Eq 6:

$$\tau_a \frac{da_j}{dt} = -a_j + \Theta\left(J^{AG} g_j - b\right) \tag{6}$$

where a_j is the activation of the jth action, g_j is the activation of the jth Go unit, b is the negative bias (i.e. threshold), Θ is a nonlinear function as in Eq 3, and τ_a is the encoding constant. J^{AG} is the weight from the jth Go unit to the jth Action unit, and was initially (i.e. before learning) randomly drawn from a Gaussian distribution with mean = 2 and s.d. = 0.2. In turn, Action nodes project excitatory connections to three distinct parts of the network simultaneously. First, Action nodes project to the cluster of excitatory neurons in the RNN representing the $i+1^{th}$ order in the sequences (blue arrow 3a in Fig 1). Second, Action nodes project to the inhibitory shared neuron (blue arrow 3b to orange node in Fig 1), that in turn globally inhibits all the clusters in the RNN. Note that the gain parameter values on Action nodes activity are larger for projections to the excitatory clusters vs inhibitory neuron of the RNN (i.e. $\gamma_E > \gamma_I$). This allows the activation of Action nodes to perturb the ratio of excitatory to inhibitory RNN input in a way that allows the ith cluster to shut down and the $i+1^{th}$ cluster to be expressed. Third, Action nodes project excitatory connections back to their corresponding No Go cells (blue arrow 3c from j^{th} Action node in the thalamus to j^{th} No Go node in the BG, see Fig 1C). The dynamics of No Go cells are in turn dictated by Eq 7:

$$\tau_n \frac{dn_j}{dt} = -n_j + J^{NA} a_j \tag{7}$$

where n_j is the activation of the jth No Go node, J^{NA} is the weight from the jth Action unit to the jth No Go unit, a_j is the activation of the jth Action node, and τ_n is the encoding constant.

In Table 1 we report the parameter values used for all 9 simulations described in the main text.

Supporting information

S1 Audio file. Simulation 7: Thunderstruck song as reproduced by the ACDC model. This audio file shows the ability of the ACDC model to learn the second guitar riff of Thunderstruck

reproduce the 16 actions (associated to 6 notes, see main text) in the correct order and tempo.
(MP4)

S2 Audio file. Simulation 7: Thunderstruck song following a bossa nova rhythm. This
audio file shows the ability of the ACDC model to perform temporal compositionality. The
ACDC model can produce flexibly produce the previously learnt guitar riff (S1 Audio file) fol-
lowing a bossa nova rhythm without any further training.
(MP4)

**S1 Video. Simulation 7: Dynamical visualization of RNN and action nodes activity coupled
with simulation-based Thunderstruck song sound.** The top left panel shows how RNN
sequential and persistent activity unfolds as a function of time. The bottom left panel is a visu-
alization of RNN dynamics as a neural trajectory in principal component (PC) space. The neu-
ral trajectory displays a pattern of sequential attractor states. The right panel displays how
activity in each Action node (and hence Thunderstruck song note) is executed at the learned
action time.
(MP4)

S2 Video. Simulation 9: Dynamical visualization of RNN and Action nodes activity. The
left panel shows how activity in each Action node is executed at the learned action time, each
color represents the activation of a specific A node in the thalamus. Given the structure and
mechanism described in Fig 1, the right panel displays the neural RNN trajectory showing that
each action execution triggers a switch from the ith to the i^{th+1} attractor state.
(MP4)

Acknowledgments

Author Contributions

Conceptualization: Cristian Buc Calderon, Tom Verguts, Michael J. Frank.

Formal analysis: Cristian Buc Calderon.

Supervision: Tom Verguts, Michael J. Frank.

Writing – original draft: Cristian Buc Calderon.

Writing – review & editing: Cristian Buc Calderon, Tom Verguts, Michael J. Frank.

References

1. Jin DZ, Fujii N, Graybiel AM. Neural representation of time in cortico-basal ganglia circuits. Proc Natl Acad Sci U S A. 2009; 106: 19156–19161. https://doi.org/10.1073/pnas.0909881106 PMID: 19850874

2. Mello GBM, Soares S, Paton JJ. A scalable population code for time in the striatum. Curr Biol. 2015; 25: 1113–1122. https://doi.org/10.1016/j.cub.2015.02.036 PMID: 25913405

3. Dhawale AK, Poddar R, Wolff SBE, Normand VA, Kopelowitz E, Ölveczky BP. Automated long-Term recording and analysis of neural activity in behaving animals. Elife. 2017; 6: 1–40. https://doi.org/10.7554/eLife.27702 PMID: 28885141

4. Gouvêa TS, Monteiro T, Motiwala A, Soares S, Machens C, Paton JJ. Striatal dynamics explain dura-tion judgments. Elife. 2015; 4: 1–14. https://doi.org/10.7554/eLife.11386 PMID: 26641377

5. Bakhurin KI, Goudar V, Shobe JL, Claar LD, Buonomano D V., Masmanidis SC. Differential Encoding of Time by Prefrontal and Striatal Network Dynamics. J Neurosci. 2017; 37: 854–870. https://doi.org/10.1523/JNEUROSCI.1789-16.2016 PMID: 28123021

6. Pastalkova E, Itskov V, Amarasingham A, Buzsaki G. Internally Generated Cell Assembly Sequences in the Rat Hippocampus. Science (80-). 2008; 321: 1322–1327. https://doi.org/10.1126/science.1159775 PMID: 18772431

7. MacDonald CJ, Tonegawa S. Crucial role for CA2 inputs in the sequential organization of CA1 time cells supporting memory. Proc Natl Acad Sci U S A. 2021; 118. https://doi.org/10.1073/pnas.2020698118 PMID: 33431691

8. Eichenbaum H. Time cells in the hippocampus: A new dimension for mapping memories. Nat Rev Neurosci. 2014; 15: 732–744. https://doi.org/10.1038/nrn3827 PMID: 25269553

9. Nicola W, Clopath C. A diversity of interneurons and Hebbian plasticity facilitate rapid compressible learning in the hippocampus. Nat Neurosci. 2019; 22: 1168–1181. https://doi.org/10.1038/s41593-019-0415-2 PMID: 31235906

10. Luczak A, Barthó P, Marguet SL, Buzsáki G, Harris KD. Sequential structure of neocortical spontaneous activity in vivo. Proc Natl Acad Sci U S A. 2007; 104: 347–352. https://doi.org/10.1073/pnas.0605643104 PMID: 17185420

11. Harvey CD, Coen P, Tank DW. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature. 2012; 484: 62–68. https://doi.org/10.1038/nature10918 PMID: 22419153

12. Remington ED, Egger SW, Narain D, Wang J, Jazayeri M. A Dynamical Systems Perspective on Flexible Motor Timing. Trends Cogn Sci. 2018; 22: 938–952. https://doi.org/10.1016/j.tics.2018.07.010 PMID: 30266152

13. Sompolinsky H, Kanter I. Temporal association in asymmetric neural networks. Phys Rev Lett. 1986; 57: 2861–2864. https://doi.org/10.1103/PhysRevLett.57.2861 PMID: 10033885

14. Kleinfeld D. Sequential state generation by model neural networks. Proc Natl Acad Sci. 1986; 83: 9469–9473. https://doi.org/10.1073/pnas.83.24.9469 PMID: 3467316

15. Herrmann M, Hertz JA, Prügel-Bennett A. Network: Computation in Neural Systems Analysis of synfire chains Analysis of synfire chains. Netw Comput Ned Syst. 1995; 6: 403–414. https://doi.org/10.1088/0954-898X

16. Graybiel AM. The Basal Ganglia and Chunking of Action Repertoires. Neurobiol Learn Mem. 1998; 70: 119–136. https://doi.org/10.1006/nlme.1998.3843 PMID: 9753592

17. Geddes CE, Li H, Jin X. Optogenetic Editing Reveals the Hierarchical Organization of Learned Action Sequences. Cell. 2018; 174: 32–43.e15. https://doi.org/10.1016/j.cell.2018.06.012 PMID: 29958111

18. Krakauer JW, Hadjiosif AM, Xu J, Wong AL, Haith AM. Motor learning. Compr Physiol. 2019; 9: 613–663. https://doi.org/10.1002/cphy.c170043 PMID: 30873583

19. Fiete IR, Senn W, Wang CZH, Hahnloser RHR. Spike-Time-Dependent Plasticity and Heterosynaptic Competition Organize Networks to Produce Long Scale-Free Sequences of Neural Activity. Neuron. 2010; 65: 563–576. https://doi.org/10.1016/j.neuron.2010.02.003 PMID: 20188660

20. Dlesmann M, Gewaltig MO, Aertsen A. Stable propagation of synchronous spiking in cortical neural networks. Nature. 1999; 402: 529–533. https://doi.org/10.1038/990101 PMID: 10591212

21. Abeles M. Corticonics: Neural circuits of the cerebral cortex. Cambridge University Press; 1991.

22. Cone I, Shouval HZ. Learning precise spatiotemporal sequences via biophysically realistic learning rules in a modular, spiking network. Elife. 2021; 10: 2020.04.17.046862. https://doi.org/10.7554/eLife.63751 PMID: 33734085

23. Pereira U, Brunel N. Unsupervised Learning of Persistent and Sequential Activity. Front Comput Neurosci. 2020; 13: 1–19. https://doi.org/10.3389/fncom.2019.00097 PMID: 32009924

24. Veliz-Cuba A, Shouval HZ, Josić K, Kilpatrick ZP. Networks that learn the precise timing of event sequences. J Comput Neurosci. 2015; 39: 235–254. https://doi.org/10.1007/s10827-015-0574-4 PMID: 26334992

25. Goodbody SJ, Wolpert DM. Temporal and amplitude generalization in motor learning. J Neurophysiol. 1998; 79: 1825–1838. https://doi.org/10.1152/jn.1998.79.4.1825 PMID: 9535951

26. Shmuelof L, Krakauer JW, Mazzoni P. How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. J Neurophysiol. 2012; 108: 578–594. https://doi.org/10.1152/jn.00856.2011 PMID: 22514286

27. Maes A, Barahona M, Clopath C. Learning spatiotemporal signals using a recurrent spiking network that discretizes time. PLoS Comput Biol. 2020; 16: 1–26. https://doi.org/10.1371/journal.pcbi.1007606 PMID: 31961853

28. Murray JM, Escola GS. Learning multiple variable-speed sequences in striatum via cortical tutoring. Elife. 2017; 6: 1–24. https://doi.org/10.7554/eLife.26084 PMID: 28481200

29. Litwin-Kumar A, Doiron B. Formation and maintenance of neuronal assemblies through synaptic plasticity. Nat Commun. 2014; 5: 1–12. https://doi.org/10.1038/ncomms6319 PMID: 25395015

30. Zenke F, Agnes EJ, Gerstner W. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. Nat Commun. 2015; 6: 1–13. https://doi.org/10.1038/ncomms7922 PMID: 25897632

31. Del Giudice P, Fusi S, Mattia M. Modelling the formation of working memory with networks of integrate-and-fire neurons connected by plastic synapses. J Physiol Paris. 2003; 97: 659–681. https://doi.org/10.1016/j.jphysparis.2004.01.021 PMID: 15242673

32. Mongillo G, Amit DJ, Brunel N. Retrospective and prospective persistent activity induced by Hebbian learning in a recurrent cortical network. Eur J Neurosci. 2003; 18: 2011–2024. https://doi.org/10.1046/j.1460-9568.2003.02908.x PMID: 14622234

33. Gillett M, Pereira U, Brunel N. Characteristics of sequential activity in networks with temporally asymmetric Hebbian learning. Proc Natl Acad Sci U S A. 2020; 117: 29948–29958. https://doi.org/10.1073/pnas.1918674117 PMID: 33177232

34. Hardy NF, Goudar V, Romero-Sosa JL, Buonomano D V. A model of temporal scaling correctly predicts that motor timing improves with speed. Nat Commun. 2018; 9: 1–14. https://doi.org/10.1038/s41467-017-02088-w PMID: 29317637

35. Orhan AE, Ma WJ. A diverse range of factors affect the nature of neural representations underlying short-term memory. Nat Neurosci. 2019; 22: 275–283. https://doi.org/10.1038/s41593-018-0314-y PMID: 30664767

36. Laje R, Buonomano D V. Robust timing and motor patterns by taming chaos in recurrent neural networks. Nat Neurosci. 2013; 16: 925–933. https://doi.org/10.1038/nn.3405 PMID: 23708144

37. Buonomano D V. Harnessing Chaos in Recurrent Neural Networks. Neuron. 2009; 63: 423–425. https://doi.org/10.1016/j.neuron.2009.08.003 PMID: 19709625

38. Rajan K, Harvey CDD, Tank DWW. Recurrent Network Models of Sequence Generation and Memory. Neuron. 2016; 90: 128–142. https://doi.org/10.1016/j.neuron.2016.02.009 PMID: 26971945

39. Goudar V, Buonomano D V. Encoding sensory and motor patterns as time-invariant trajectories in recurrent neural networks. Elife. 2018; 7: 1–28. https://doi.org/10.7554/eLife.31134 PMID: 29537963

40. Logiaco L, Abbott LF, Escola S. Thalamic control of cortical dynamics in a model of flexible motor sequencing. Cell Rep. 2021; 35: 109090. https://doi.org/10.1016/j.celrep.2021.109090 PMID: 34077721

41. Sussillo D, Abbott LF. Generating Coherent Patterns of Activity from Chaotic Neural Networks. Neuron. 2009; 63: 544–557. https://doi.org/10.1016/j.neuron.2009.07.018 PMID: 19709635

42. Bellec G, Scherr F, Subramoney A, Hajek E, Salaj D, Legenstein R, et al. A solution to the learning dilemma for recurrent networks of spiking neurons. Nat Commun. 2020; 11: 1–15. https://doi.org/10.1038/s41467-019-13993-7 PMID: 31911652

43. Miconi T. Biologically plausible learning in recurrent neural networks reproduces neural dynamics observed during cognitive tasks. Elife. 2017; 6: 1–24. https://doi.org/10.7554/eLife.20899 PMID: 28230528

44. Murray JM. Local online learning in recurrent networks with random feedback. Elife. 2019; 8: 1–25. https://doi.org/10.7554/eLife.43299 PMID: 31124785

45. Wang J, Narain D, Hosseini EA, Jazayeri M. Flexible timing by temporal scaling of cortical responses. Nat Neurosci. 2018; 21: 102–112. https://doi.org/10.1038/s41593-017-0028-6 PMID: 29203897

46. Recanatesi S, Pereira-Obilinovic U, Murakami M, Mainen Z, Mazzucato L. Metastable attractors explain the variable timing of stable behavioral action sequences. Neuron. 2021; 1–15. https://doi.org/10.1016/j.neuron.2020.12.009 PMID: 33412092

47. Franklin NT, Frank MJ. Compositional clustering in task structure learning. Daunizeau J, editor. PLOS Comput Biol. 2018; 14: e1006116. https://doi.org/10.1371/journal.pcbi.1006116 PMID: 29672581

48. Franklin NT, Frank MJ. Generalizing to generalize: Humans flexibly switch between compositional and conjunctive structures during reinforcement learning. PLoS Computational Biology. 2020. https://doi.org/10.1371/journal.pcbi.1007720 PMID: 32282795

49. Burke DA, Rotstein HG, Alvarez VA. Striatal Local Circuitry: A New Framework for Lateral Inhibition. Neuron. 2017; 96: 267–284. https://doi.org/10.1016/j.neuron.2017.09.019 PMID: 29024654

50. Halassa MM, Sherman SM. Thalamocortical Circuit Motifs: A General Framework. Neuron. 2019; 103: 762–770. https://doi.org/10.1016/j.neuron.2019.06.005 PMID: 31487527

51. Rikhye R V., Wimmer RD, Halassa MM. Toward an Integrative Theory of Thalamic Function. Annu Rev Neurosci. 2018; 41: 163–183. https://doi.org/10.1146/annurev-neuro-080317-062144 PMID: 29618284

52. Rikhye R V., Gilra A, Halassa MM. Thalamic regulation of switching between cortical representations enables cognitive flexibility. Nat Neurosci. 2018; 21: 1753–1763. https://doi.org/10.1038/s41593-018-0269-z PMID: 30455456

53. Wymbs NF, Bassett DS, Mucha PJ, Porter MA, Grafton ST. Article Differential Recruitment of the Sensorimotor Putamen and Frontoparietal Cortex during Motor Chunking in Humans. Neuron. 2012; 74: 936–946. https://doi.org/10.1016/j.neuron.2012.03.038 PMID: 22681696

54. Lungu O, Monchi O, Albouy G, Jubault T, Ballarin E, Burnod Y, et al. Striatal and hippocampal involvement in motor sequence chunking depends on the learning strategy. PLoS One. 2014; 9: 25–27. https://doi.org/10.1371/journal.pone.0103885 PMID: 25148078

55. Doyon J, Gabitov E, Vahdat S, Lungu O, Boutin A. ScienceDirect Current issues related to motor sequence learning in humans. Curr Opin Behav Sci. 2017; 20: 89–97. https://doi.org/10.1016/j.cobeha.2017.11.012

56. Graybiel AM, Grafton ST. The striatum: Where skills and habits meet. Cold Spring Harb Perspect Biol. 2015; 7: 1–14. https://doi.org/10.1101/cshperspect.a021691 PMID: 26238359

57. Boutin A, Massen C, Heuer H. Modality-specific organization in the representation of sensorimotor sequences. Front Psychol. 2013; 4: 1–9. https://doi.org/10.3389/fpsyg.2013.00001 PMID: 23382719

58. Abrahamse EL, Ruitenberg MFL, de Kleine E, Verwey WB. Control of automated behavior: Insights from the discrete sequence production task. Front Hum Neurosci. 2013; 7: 1–16. https://doi.org/10.3389/fnhum.2013.00001 PMID: 23355817

59. Diedrichsen J, Kornysheva K. Motor skill learning between selection and execution. Trends Cogn Sci. 2015; 19: 227–233. https://doi.org/10.1016/j.tics.2015.02.003 PMID: 25746123

60. Verwey WB, Shea CH, Wright DL. A cognitive framework for explaining serial processing and sequence execution strategies. Psychon Bull Rev. 2014; 22: 54–77. https://doi.org/10.3758/s13423-014-0773-4 PMID: 25421407

61. Kornysheva K, Bush D, Meyer SS, Sadnicka A, Barnes G, Burgess N. Neural Competitive Queuing of Ordinal Structure Underlies Skilled Sequential Action. Neuron. 2019; 101: 1166–1180.e3. https://doi.org/10.1016/j.neuron.2019.01.018 PMID: 30744987

62. Forstmann BU, Dutilh G, Brown SD, Neumann J, von Cramon DY, Ridderinkhof KR, et al. Striatum and pre-SMA facilitate decision-making under time pressure. Proc Natl Acad Sci U S A. 2008; 105: 17538–17542. https://doi.org/10.1073/pnas.0805903105 PMID: 18981414

63. Acerbi L, Wolpert DM, Vijayakumar S. Internal Representations of Temporal Statistics and Feedback Calibrate Motor-Sensory Interval Timing. PLoS Comput Biol. 2012; 8. https://doi.org/10.1371/journal.pcbi.1002771 PMID: 23209386

64. Rakitin BC, Penney TB, Gibbon J, Malapani C, Hinton SC, Meck WH. Scalar expectancy theory and peak-interval timing in humans. J Exp Psychol Anim Behav Process. 1998; 24: 15–33. https://doi.org/10.1037//0097-7403.24.1.15 PMID: 9438963

65. Jazayeri M, Shadlen MN. Temporal context calibrates interval timing. Nat Neurosci. 2010; 13: 1020–1026. https://doi.org/10.1038/nn.2590 PMID: 20581842

66. Ivry RB, Hazeltine RE. Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. J Exp Psychol Hum Percept Perform. 1995; 21: 3–18. https://doi.org/10.1037//0096-1523.21.1.3 PMID: 7707031

67. Egger SW, Le NM, Jazayeri M. A neural circuit model for human sensorimotor timing. Nat Commun. 2020; 11: 1–14. https://doi.org/10.1038/s41467-019-13993-7 PMID: 31911652

68. Ratcliff R, Rouder JN. Modelling response times for two choice decisions. Psychol Sci. 1998; 9: 347–356.

69. Miyashita Y, Chang HS. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. Nature. 1988; 331: 68–70. https://doi.org/10.1038/331068a0 PMID: 3340148

70. Nakamura K, Kubota K. Mnemonic firing of neurons in the monkey temporal pole during a visual recognition memory task. J Neurophysiol. 1995; 74: 162–178. https://doi.org/10.1152/jn.1995.74.1.162 PMID: 7472321

71. Erickson CA, Desimone R. Responses of macaque perirhinal neurons during and after visual stimulus association learning. J Neurosci. 1999; 19: 10404–10416. https://doi.org/10.1523/JNEUROSCI.19-23-10404.1999 PMID: 10575038

72. Koch KW, Fuster JM. Unit activity in monkey parietal cortex related to haptic perception and temporary memory. Exp Brain Res. 1989. https://doi.org/10.1007/BF00247889 PMID: 2767186

73. Chafee M V., Goldman-Rakic PS. Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. J Neurophysiol. 1998; 79: 2919–2940. https://doi.org/10.1152/jn.1998.79.6.2919 PMID: 9636098

74. Gail A, Andersen R a. Neural dynamics in monkey parietal reach region reflect context-specific sensorimotor transformations. J Neurosci. 2006; 26: 9376–9384. https://doi.org/10.1523/JNEUROSCI.1570-06.2006 PMID: 16971521

75. Klaes C, Westendorff S, Chakrabarti S, Gail A. Choosing goals, not rules: deciding among rule-based action plans. Neuron. 2011; 70: 536–48. https://doi.org/10.1016/j.neuron.2011.02.053 PMID: 21555078

76. Cisek P, Kalaska JF. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. Neuron. 2005; 45: 801–14. https://doi.org/10.1016/j.neuron.2005.01.027 PMID: 15748854

77. Funahashi S, Bruce CJ, Goldman-Rakic PS. Visuospatial coding in primate prefrontal neurons revealed by oculomotor paradigms. J Neurophysiol. 1990; 63: 814–831. https://doi.org/10.1152/jn.1990.63.4.814 PMID: 2341879

78. Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol. 1989; 61: 331–349. https://doi.org/10.1152/jn.1989.61.2.331 PMID: 2918358

79. Miller EK, Erickson CA, Desimone R. Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J Neurosci. 1996; 16: 5154–5167. https://doi.org/10.1523/JNEUROSCI.16-16-05154.1996 PMID: 8756444

80. Wang XJ. Synaptic reverberation underlying mnemonic persistent activity. Trends Neurosci. 2001; 24: 455–463. https://doi.org/10.1016/s0166-2236(00)01868-3 PMID: 11476885

81. Durstewitz D, Seamans JK, Sejnowski TJ. Neurocomputational Models of Working Memory. Nat Neurosci. 2000; 3: 1184–1191. https://doi.org/10.1038/81460 PMID: 11127836

82. Brunel N. Dynamics and Plasticity of Stimulus-selective Persistent Activity in Cortical Network Models. Cereb Cortex. 2003; 13: 1151–1161. https://doi.org/10.1093/cercor/bhg096 PMID: 14576207

83. Hahnloser RHR, Kozhevnikov AA, Fee MS. An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature. 2002; 419: 65–70. https://doi.org/10.1038/nature00974 PMID: 12214232

84. Okubo TS, Mackevicius EL, Payne HL, Lynch GF, Fee MS. Growth and splitting of neural sequences in songbird vocal development. Nature. 2015; 528: 352–357. https://doi.org/10.1038/nature15741 PMID: 26618871

85. Kozhevnikov AA, Fee MS. Singing-related activity of identified HVC neurons in the zebra finch. J Neurophysiol. 2007; 97: 4271–4283. https://doi.org/10.1152/jn.00952.2006 PMID: 17182906

86. Amador A, Perl YS, Mindlin GB, Margoliash D. Elemental gesture dynamics are encoded by song premotor cortical neurons. Nature. 2013; 495: 59–64. https://doi.org/10.1038/nature11967 PMID: 23446354

87. Recanatesi S, Pereira U, Murakami M, Mainen Z, Mazzucato L. Metastable attractors explain the variable timing of stable behavioral action sequences. bioRxiv. 2020; 0–40. https://doi.org/10.1101/2020.01.24.919217

88. Rueda-Orozco PE, Robbe D. The striatum multiplexes contextual and kinematic information to constrain motor habits execution. Nat Neurosci. 2015; 18: 453–462. https://doi.org/10.1038/nn.3924 PMID: 25622144

89. Verguts T, Vassena E, Silvetti M. Adaptive effort investment in cognitive and physical tasks: a neurocomputational model. Front Behav Neurosci. 2015; 9: 57. https://doi.org/10.3389/fnbeh.2015.00057 PMID: 25805978

90. Collins AGE, Frank MJ. Cognitive control over learning: creating, clustering, and generalizing task-set structure. Psychol Rev. 2013; 120: 190–229. https://doi.org/10.1037/a0030852 PMID: 23356780

91. Calderon CB, Gevers W, Verguts T. The unfolding action model of initiation times, movement times, and movement paths. Psychol Rev. 2018; 1–65. https://doi.org/10.1037/rev0000074 PMID: 29035078

92. Ju H, Bassett DS. Dynamic representations in networked neural systems. Nat Neurosci. 2020; 23: 908–917. https://doi.org/10.1038/s41593-020-0653-3 PMID: 32541963

93. Burgess N, Hitch GJ. A revised model of short-term memory and long-term learning of verbal sequences. J Mem Lang. 2006; 55: 627–652. https://doi.org/10.1016/j.jml.2006.08.005

94. Burgess N, Hitch GJ. Memory for serial order: A network model of the phonological loop and its timing. Psychol Rev. 1999; 106: 551–581. https://doi.org/10.1037/0033-295X.106.3.551

95. Page MPA, Norris D. The primacy model: A new model of immediate serial recall. Psychol Rev. 1998; 105: 761–781. https://doi.org/10.1037/0033-295x.105.4.761-781 PMID: 9830378

96. Hartley T, Houghton G. A linguistically constrained model of short-term memory for nonwords. J Mem Lang. 1996; 35: 1–31. https://doi.org/10.1006/jmla.1996.0001

97. Kawai R, Markman T, Poddar R, Ko R, Fantana AL, Dhawale AK, et al. Motor Cortex Is Required for Learning but Not for Executing a Motor Skill. Neuron. 2015; 86: 800–812. https://doi.org/10.1016/j.neuron.2015.03.024 PMID: 25892304

98. Otchy TM, Wolff SBE, Rhee JY, Pehlevan C, Kawai R, Kempf A, et al. Acute off-target effects of neural circuit manipulations. Nature. 2015; 528: 358–363. https://doi.org/10.1038/nature16442 PMID: 26649821

99. Koziol LF, Budding DE. Subcortical structures and cognition: Implications for neuropsychological assessment. Springer Science & Business Media.; 2009.

100. Collins AGE, Frank MJ. Motor Demands Constrain Cognitive Rule Structures. PLoS Comput Biol. 2016; 12: 1–17. https://doi.org/10.1371/journal.pcbi.1004785 PMID: 26966909

101. Kozachkov L, Michmizos KP. Sequence Learning in Associative Neuronal-Astrocytic Networks. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2020. pp. 349–360. https://doi.org/10.1371/journal.pcbi.1007659 PMID: 32764745

102. Florian R V. Reinforcement learning through modulation of spike-timing-dependent synaptic plasticity. Neural Comput. 2007; 19: 1468–1502. https://doi.org/10.1162/neco.2007.19.6.1468 PMID: 17444757

103. Izhikevich EM. Solving the distal reward problem through linkage of STDP and dopamine signaling. Cereb Cortex. 2007; 17: 2443–2452. https://doi.org/10.1093/cercor/bhl152 PMID: 17220510

104. Frémaux N, Sprekeler H, Gerstner W. Functional requirements for reward-modulated spike-timing-dependent plasticity. J Neurosci. 2010; 30: 13326–13337. https://doi.org/10.1523/JNEUROSCI.6249-09.2010 PMID: 20926659

105. Soltoggio A, Steil JJ. Solving the distal reward problem with rare correlations. Neural Comput. 2013; 25: 940–978. https://doi.org/10.1162/NECO_a_00419 PMID: 23339615

106. Bellec G, Scherr F, Hajek E, Salaj D, Legenstein R, Maass W. Biologically inspired alternatives to backpropagation through time for learning in recurrent neural nets. arXiv. 2019; 1–37.

107. Sussillo D, Churchland MM, Kaufman MT, Shenoy K V. A neural network that finds a naturalistic solution for the production of muscle activity. Nat Neurosci. 2015; 18: 1025–1033. https://doi.org/10.1038/nn.4042 PMID: 26075643

108. Russo AA, Bittner SR, Perkins SM, Seely JS, London BM, Lara AH, et al. Motor Cortex Embeds Muscle-like Commands in an Untangled Population Response. Neuron. 2018; 97: 953–966.e8. https://doi.org/10.1016/j.neuron.2018.01.004 PMID: 29398358

109. Watabe-Uchida M, Eshel N, Uchida N. Neural Circuitry of Reward Prediction Error. Annu Rev Neurosci. 2017; 40: 373–394. https://doi.org/10.1146/annurev-neuro-072116-031109 PMID: 28441114

110. Beierholm U, Guitart-Masip M, Economides M, Chowdhury R, Düzel E, Dolan R, et al. Dopamine modulates reward-related vigor. Neuropsychopharmacology. 2013; 38: 1495–1503. https://doi.org/10.1038/npp.2013.48 PMID: 23419875

111. Panigrahi B, Martin KA, Li Y, Graves AR, Vollmer A, Olson L, et al. Dopamine Is Required for the Neural Representation and Control of Movement Vigor. Cell. 2015; 162: 1418–1430. https://doi.org/10.1016/j.cell.2015.08.014 PMID: 26359992

112. Zénon A, Devesse S, Olivier E. Dopamine manipulation affects response vigor independently of opportunity cost. J Neurosci. 2016; 36: 9516–9525. https://doi.org/10.1523/JNEUROSCI.4467-15.2016 PMID: 27629704

113. Berke JD. What does dopamine mean? Nat Neurosci. 2018; 21: 787–793. https://doi.org/10.1038/s41593-018-0152-y PMID: 29760524

114. Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, et al. Mesolimbic dopamine signals the value of work. Nat Neurosci. 2015; 19: 117–126. https://doi.org/10.1038/nn.4173 PMID: 26595651

115. Hamid AA, Frank MJ, Moore CI. Dopamine waves as a mechanism for spatiotemporal credit assignment. Cell. 2021. https://doi.org/10.1016/j.cell.2021.03.046 PMID: 33861952

116. Gaidica M, Hurst A, Cyr C, Leventhal DK. Distinct populations of motor thalamic neurons encode action initiation, action selection, and movement vigor. J Neurosci. 2018; 38: 6563–6573. https://doi.org/10.1523/JNEUROSCI.0463-18.2018 PMID: 29934350

117. Sedaghat-Nejad E, Herzfeld DJ, Shadmehr R. Reward prediction error modulates saccade vigor. J Neurosci. 2019; 39: 5010–5017. https://doi.org/10.1523/JNEUROSCI.0432-19.2019 PMID: 31015343

118.   Augustin SM, Loewinger GC, O'Neal TJ, Kravitz A V., Lovinger DM. Dopamine D2 receptor signaling on iMSNs is required for initiation and vigor of learned actions. Neuropsychopharmacology. 2020; 45: 2087–2097. https://doi.org/10.1038/s41386-020-00799-1 PMID: 32811899

119.   van Gaalen MM, van Koten R, Schoffelmeer ANM, Vanderschuren LJMJ. Critical Involvement of Dopaminergic Neurotransmission in Impulsive Decision Making. Biol Psychiatry. 2006; 60: 66–73. https://doi.org/10.1016/j.biopsych.2005.06.005 PMID: 16125144

120.   Pattij T, Vanderschuren LJMJ. The neuropharmacology of impulsive behaviour. Trends Pharmacol Sci. 2008; 29: 192–199. https://doi.org/10.1016/j.tips.2008.01.002 PMID: 18304658

121.   Buckholtz JW, Treadway MT, Cowan RL, Woodward ND, Li R, Ansari MS, et al. Dopaminergic network differences in human impulsivity. Science (80-). 2010; 329: 532. https://doi.org/10.1126/science.1185778 PMID: 20671181

122.   Pine A, Shiner T, Seymour B, Dolan RJ. Dopamine, time, and impulsivity in humans. J Neurosci. 2010; 30: 8888–8896. https://doi.org/10.1523/JNEUROSCI.6028-09.2010 PMID: 20592211

123.   Dalley JW, Roiser JP. Dopamine, serotonin and impulsivity. Neuroscience. 2012; 215: 42–58. https://doi.org/10.1016/j.neuroscience.2012.03.065 PMID: 22542672

124.   Economidou D, Theobald DEH, Robbins TW, Everitt BJ, Dalley JW. Norepinephrine and dopamine modulate impulsivity on the five-choice serial reaction time task through opponent actions in the shell and core sub-regions of the nucleus accumbens. Neuropsychopharmacology. 2012; 37: 2057–2066. https://doi.org/10.1038/npp.2012.53 PMID: 22510726

125.   Frank MJ, Samanta J, Moustafa AA, Sherman SJ. Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. Science. 2007; 318: 1309–12. https://doi.org/10.1126/science.1146157 PMID: 17962524

126.   Lake JI, Meck WH. Differential effects of amphetamine and haloperidol on temporal reproduction: Dopaminergic regulation of attention and clock speed. Neuropsychologia. 2013; 51: 284–292. https://doi.org/10.1016/j.neuropsychologia.2012.09.014 PMID: 22982605

127.   Ratcliff R. A theory of memory retrieval. Psychol Rev. 1978; 85: 59–108. https://doi.org/10.1037/0033-295X.85.2.59

128.   Beste C, Adelhöfer N, Gohil K, Passow S, Roessner V, Li SC. Dopamine modulates the efficiency of sensory evidence accumulation during perceptual decision making. Int J Neuropsychopharmacol. 2018; 21: 649–655. https://doi.org/10.1093/ijnp/pyy019 PMID: 29618012

129.   Yousif N, Fu RZ, Abou-El-Ela Bourquin B, Bhrugubanda V, Schultz SR, Seemungal BM. Dopamine activation preserves visual motion perception despite noise interference of human v5/mt. J Neurosci. 2016; 36: 9303–9312. https://doi.org/10.1523/JNEUROSCI.4452-15.2016 PMID: 27605607

130.   Westbrook A, van den Bosch R, Määttä JI, Hofmans L, Papadopetraki D, Cools R, et al. Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work. Science (80-). 2020; 367: 1362–1366. https://doi.org/10.1126/science.aaz5891 PMID: 32193325

131.   Heitz RP. The speed-accuracy tradeoff: History, physiology, methodology, and behavior. Front Neurosci. 2014; 8: 1–19. https://doi.org/10.3389/fnins.2014.00001 PMID: 24478622

132.   Wiecki TV, Frank MJ. A computational model of inhibitory control in frontal cortex and basal ganglia. Psychol Rev. 2013; 120: 329–355. https://doi.org/10.1037/a0031542 PMID: 23586447

133.   Lloyd K, Dayan P. Tamping Ramping: Algorithmic, Implementational, and Computational Explanations of Phasic Dopamine Signals in the Accumbens. PLoS Comput Biol. 2015; 11: 1–34. https://doi.org/10.1371/journal.pcbi.1004622 PMID: 26699940

134.   Frank MJ. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. Neural Netw. 2006; 19: 1120–36. https://doi.org/10.1016/j.neunet.2006.03.006 PMID: 16945502

135.   Cavanagh JF, Wiecki TV, Cohen MX, Figueroa CM, Samanta J, Sherman SJ, et al. Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nat Neurosci. 2011; 14: 1462–7. https://doi.org/10.1038/nn.2925 PMID: 21946325

136.   Herz DM, Zavala BA, Bogacz R, Brown P. Neural Correlates of Decision Thresholds in the Human Subthalamic Nucleus. Curr Biol. 2016; 26: 916–920. https://doi.org/10.1016/j.cub.2016.01.051 PMID: 26996501

137.   Montague PR, Dayan P, Person C, Sejnowski TJ. Bee foraging in uncertain environments. Nature. 1995. pp. 725–728. http://papers.cnl.salk.edu/PDFs/Bee%20Foraging%20in%20Uncertain%20Environments%20Using%20Predictive%20Hebbian%20Learning%201995-3013.pdf https://doi.org/10.1038/377725a0 PMID: 7477260

138.   Hoerzer GM, Legenstein R, Maass W. Emergence of complex computational structures from chaotic neural networks through reward-modulated hebbian learning. Cereb Cortex. 2014; 24: 677–690. https://doi.org/10.1093/cercor/bhs348 PMID: 23146969

139. Bailey CH, Giustetto M, Huang YY, Hawkins RD, Kandel ER. Is Heterosynaptic modulation essential for stabilizing hebbian plasiticity and memory. Nat Rev Neurosci. 2000; 1: 11–20. https://doi.org/10.1038/35036191 PMID: 11252764

140. Frank MJ. Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J Cogn Neurosci. 2005; 17: 51–72. https://doi.org/10.1162/0898929052880093 PMID: 15701239

141. Gerstner W, Lehmann M, Liakoni V, Corneil D, Brea J. Eligibility Traces and Plasticity on Behavioral Time Scales: Experimental Support of NeoHebbian Three-Factor Learning Rules. Front Neural Circuits. 2018; 12: 1–16. https://doi.org/10.3389/fncir.2018.00001 PMID: 29403360

142. Doya K. Complementary roles of basal ganglia and cerebellum in learning and motor control. Curr Opin Neurobiol. 2000; 10: 732–739. https://doi.org/10.1016/s0959-4388(00)00153-7 PMID: 11240282

143. Legenstein R, Pecevski D, Maass W. A learning theory for reward-modulated spike-timing-dependent plasticity with application to biofeedback. PLoS Comput Biol. 2008; 4. https://doi.org/10.1371/journal.pcbi.1000180 PMID: 18846203

144. Lalive AL, Lien AD, Roseberry TK, Donahue CH, Kreitzer AC. Motor thalamus supports striatum-driven reinforcement. Elife. 2018; 7: 1–22. https://doi.org/10.7554/eLife.34032 PMID: 30295606

145. Mcclure SM, Berns GS, Montague PR. Temporal Prediction Errors in a Passive Learning Task Activate Human Striatum. 2003; 38: 339–346.

146. Doherty JO, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. 2004; 304: 452–455.

147. Badre D, Frank MJ. Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: evidence from fMRI. Cereb Cortex. 2012; 22: 527–36. https://doi.org/10.1093/cercor/bhr117 PMID: 21693491

148. Lillicrap TP, Cownden D, Tweed DB, Akerman CJ. Random synaptic feedback weights support error backpropagation for deep learning. Nat Commun. 2016; 7: 1–10. https://doi.org/10.1038/ncomms13276 PMID: 27824044

149. O'Reilly RC, Munakata Y, Frank MJ, Hazy TE. Computational cognitive neuroscience. PediaPress; 2012.

150. Perich MG, Rajan K. Rethinking brain-wide interactions through multi-region 'network of networks' models. Curr Opin Neurobiol. 2020; 65: 146–151. https://doi.org/10.1016/j.conb.2020.11.003 PMID: 33254073

151. Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W. Inhibitory Plasticity Balances Excitation and Inhibition in Sensory Pathways and Memory Networks. Science (80-). 2011; 334: 1569–1573. https://doi.org/10.1126/science.1211095 PMID: 22075724

152. Fusi S, Miller EK, Rigotti M. Why neurons mix: High dimensionality for higher cognition. Curr Opin Neurobiol. 2016; 37: 66–74. https://doi.org/10.1016/j.conb.2016.01.010 PMID: 26851755

153. Badre D, Bhandari A, Keglovits H, Kikumoto A. The dimensionality of neural representations for control. Current Opinion in Behavioral Sciences. Elsevier Ltd; 2021. pp. 20–28. https://doi.org/10.1016/j.cobeha.2020.07.002 PMID: 32864401

154. Márton CD, Schultz SR, Averbeck BB. Learning to select actions shapes recurrent dynamics in the corticostriatal system. Neural Networks. 2020; 132: 375–393. https://doi.org/10.1016/j.neunet.2020.09.008 PMID: 32992244

155. Erlhagen W, Schöner G. Dynamic field theory of movement preparation. Psychol Rev. 2002; 109: 545–572. https://doi.org/10.1037/0033-295x.109.3.545 PMID: 12088245

156. Klaes C, Schneegans S, Schöner G, Gail A. Sensorimotor Learning Biases Choice Behavior: A Learning Neural Field Model for Decision Making. PLoS Comput Biol. 2012; 8: e1002774. https://doi.org/10.1371/journal.pcbi.1002774 PMID: 23166483

157. Cisek P. Integrated neural processes for defining potential actions and deciding between them: a computational model. J Neurosci. 2006; 26: 9761–70. https://doi.org/10.1523/JNEUROSCI.5605-05.2006 PMID: 16988047

158. Christopoulos V, Bonaiuto J, Andersen R a. A Biologically Plausible Computational Theory for Value Integration and Action Selection in Decisions with Competing Alternatives. PLOS Comput Biol. 2015; 11: e1004104. https://doi.org/10.1371/journal.pcbi.1004104 PMID: 25803729

159. Thach WT, Goodkin HP, Keating JG. The cerebellum and the adaptive coordination of movement. Annual Review of Neuroscience. 1992. pp. 403–442. https://doi.org/10.1146/annurev.ne.15.030192.002155 PMID: 1575449

160. Bastian AJ. Learning to predict the future: the cerebellum adapts feedforward movement control. Current Opinion in Neurobiology. 2006. pp. 645–649. https://doi.org/10.1016/j.conb.2006.08.016 PMID: 17071073

161. Yang GR, Joglekar MR, Song HF, Newsome WT, Wang XJ. Task representations in neural networks trained to perform many cognitive tasks. Nat Neurosci. 2019; 22: 297–306. https://doi.org/10.1038/s41593-018-0310-2 PMID: 30643294

162. Baram AB, Muller TH, Nili H, Garvert MM, Behrens TEJ. Entorhinal and ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning problems. Neuron. 2021; 109: 713–723.e7. https://doi.org/10.1016/j.neuron.2020.11.024 PMID: 33357385

163. Whittington JCR, Muller TH, Mark S, Chen G, Barry C, Burgess N, et al. The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through Generalization in the Hippocampal Formation. Cell. 2020; 183: 1249–1263.e23. https://doi.org/10.1016/j.cell.2020.10.024 PMID: 33181068

164. Collins AGE, Frank MJ. Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. Cognition. 2016; 152: 160–169. https://doi.org/10.1016/j.cognition.2016.04.002 PMID: 27082659

165. Lehnert L, Littman ML, Frank MJ. Reward-predictive representations generalize across tasks in reinforcement learning. PLoS Comput Biol. 2020; 16: 1–27. https://doi.org/10.1371/journal.pcbi.1008317 PMID: 33057329

166. Amit DJ. Neural networks counting chimes. Proc Natl Acad Sci U S A. 1988; 85: 2141–2145. https://doi.org/10.1073/pnas.85.7.2141 PMID: 3353371

167. O'Reilly RC, Frank MJ. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. Neural Comput. 2006; 18: 283–328. https://doi.org/10.1162/089976606775093909 PMID: 16378516

168. Ratcliff R, Frank MJ. Reinforcement-based decision making in corticostriatal circuits: Mutual constraints by neurocomputational and diffusion models. Neural Comput. 2012; 24: 1186–1229. https://doi.org/10.1162/NECO_a_00270 PMID: 22295983

169. Doi T, Fan Y, Gold JI, Ding L. The caudate nucleus contributes causally to decisions that balance reward and uncertain visual information. Elife. 2020; 9: 568733. https://doi.org/10.7554/eLife.56694 PMID: 32568068

170. Mink JW. The basal ganglia: Focused selection and inhibition of competing motor programs. Prog Neurobiol. 1996; 50: 381–425. https://doi.org/10.1016/s0301-0082(96)00042-1 PMID: 9004351

171. Alexander GE, Crutcher MD. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. Trends in Neurosciences. 1990. pp. 266–271. https://doi.org/10.1016/0166-2236(90)90107-l PMID: 1695401

172. Franklin NT, Frank MJ. A cholinergic feedback circuit to regulate striatal population uncertainty and optimize reinforcement learning. Elife. 2015; 4: 1–29. https://doi.org/10.7554/eLife.12029 PMID: 26705698

173. Kim J, Kim Y, Nakajima R, Shin A, Jeong M, Park AH, et al. Inhibitory Basal Ganglia Inputs Induce Excitatory Motor Signals in the Thalamus. Neuron. 2017; 95: 1181–1196.e8. https://doi.org/10.1016/j.neuron.2017.08.028 PMID: 28858620

174. Lo CC, Wang XJ. Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. Nat Neurosci. 2006; 9: 956–963. https://doi.org/10.1038/nn1722 PMID: 16767089

175. Schmitt LI, Wimmer RD, Nakajima M, Happ M, Mofakham S, Halassa MM. Thalamic amplification of cortical connectivity sustains attentional control. Nature. 2017; 545: 219–223. https://doi.org/10.1038/nature22073 PMID: 28467827

176. Taverna S, Ilijic E, Surmeier DJ. Recurrent collateral connections of striatal medium spiny neurons are disrupted in models of Parkinson's disease. J Neurosci. 2008; 28: 5504–5512. https://doi.org/10.1523/JNEUROSCI.5493-07.2008 PMID: 18495884

177. Dobbs LKK, Kaplan ARR, Lemos JCC, Matsui A, Rubinstein M, Alvarez VAA. Dopamine Regulation of Lateral Inhibition between Striatal Neurons Gates the Stimulant Actions of Cocaine. Neuron. 2016; 90: 1100–1113. https://doi.org/10.1016/j.neuron.2016.04.031 PMID: 27181061

178. Averbeck BB, Sohn JW, Lee D. Activity in prefrontal cortex during dynamic selection of action sequences. Nat Neurosci. 2006; 9: 276–282. https://doi.org/10.1038/nn1634 PMID: 16429134

179. Berdyyeva TK, Olson CR. Monkey Supplementary Eye Field Neurons Signal the Ordinal Position of Both Actions and Objects. J Neurosci. 2009; 29: 591–599. https://doi.org/10.1523/JNEUROSCI.4803-08.2009 PMID: 19158286

180. Isoda M, Tanji J. Contrasting Neuronal Activity in the Supplementary and Frontal Eye Fields during Temporal Organization of Multiple Saccades. J Neurophysiol. 2003; 90: 3054–3065. https://doi.org/10.1152/jn.00367.2003 PMID: 12904333

181. Isoda M, Tanji J. Participation of the primate presupplementary motor area in sequencing multiple saccades. J Neurophysiol. 2004; 92: 653–659. https://doi.org/10.1152/jn.01201.2003 PMID: 14985413

182. Shima K, Tanji J. Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. J Neurophysiol. 2000; 84: 2148–60. 80:3247–3260 https://doi.org/10.1152/jn.2000.84.4.2148 PMID: 11024102

183. Clower WT, Alexander GE. Movement sequence-related activity reflecting numerical order of components in supplementary and presupplementary motor areas. J Neurophysiol. 1998; 80: 1562–1566. https://doi.org/10.1152/jn.1998.80.3.1562 PMID: 9744961

184. Salinas E. Rank-Order-Selective Neurons Form a Temporal Basis Set for the Generation of Motor Sequences. J Neurosci. 2009; 29: 4369–4380. https://doi.org/10.1523/JNEUROSCI.0164-09.2009 PMID: 19357265

185. Schwartze M, Keller PE, Patel AD, Kotz SA. The impact of basal ganglia lesions on sensorimotor synchronization, spontaneous motor tempo, and the detection of tempo changes. Behav Brain Res. 2011; 216: 685–691. https://doi.org/10.1016/j.bbr.2010.09.015 PMID: 20883725

186. Gershman SJ, Moustafa AA, Ludvig EA. Time representation in reinforcement learning models of the basal ganglia. Front Comput Neurosci. 2014; 7: 1–8. https://doi.org/10.3389/fncom.2013.00194 PMID: 24409138

187. Jones CRG, Jahanshahi M. Contributions of the Basal Ganglia to Temporal Processing: Evidence from Parkinson's Disease. Timing Time Percept. 2014; 2: 87–127. https://doi.org/10.1163/22134468-00002009

188. Thura D, Cisek P. The Basal Ganglia Do Not Select Reach Targets but Control the Urgency of Commitment. Neuron. 2017; 95: 1160–1170.e5. https://doi.org/10.1016/j.neuron.2017.07.039 PMID: 28823728

189. Paton JJ, Buonomano D V. The Neural Basis of Timing: Distributed Mechanisms for Diverse Functions. Neuron. 2018; 98: 687–705. https://doi.org/10.1016/j.neuron.2018.03.045 PMID: 29772201

190. Gurney K, Prescott TJ, Redgrave P. A computational model of action selection in the basal ganglia. I. A new functional anatomy. Biol Cybern. 2001; 84: 401–410. https://doi.org/10.1007/PL00007984 PMID: 11417052

191. Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci. 1986; VOL. 9: 357–381. https://doi.org/10.1146/annurev.ne.09.030186.002041 PMID: 3085570

192. Oh SW, Harris JA, Ng L, Winslow B, Cain N, Mihalas S, et al. A mesoscale connectome of the mouse brain. Nature. 2014; 508: 207–214. https://doi.org/10.1038/nature13186 PMID: 24695228

193. Hintiryan H, Foster NN, Bowman I, Bay M, Song MY, Gou L, et al. The mouse cortico-striatal projectome. Nat Neurosci. 2016; 19: 1100–1114. https://doi.org/10.1038/nn.4332 PMID: 27322419

194. Hunnicutt BJ, Jongbloets BC, Birdsong WT, Gertz KJ, Zhong H, Mao T. A comprehensive excitatory input map of the striatum reveals novel functional organization. Elife. 2016; 5: 1–32. https://doi.org/10.7554/eLife.19103 PMID: 27892854

195. Peters AJ, Fabre JMJ, Steinmetz NA, Harris KD, Carandini M. Striatal activity topographically reflects cortical activity. Nature. 2021; 591. https://doi.org/10.1038/s41586-020-03166-8 PMID: 33473213

196. Gerardin E, Lehéricy S, Pochon JB, Du Montcel ST, Mangin JF, Poupon F, et al. Foot, hand, face and eye representation in the human striatum. Cereb Cortex. 2003; 13: 162–169. https://doi.org/10.1093/cercor/13.2.162 PMID: 12507947

197. McHaffie JG, Stanford TR, Stein BE, Coizet V, Redgrave P. Subcortical loops through the basal ganglia. Trends Neurosci. 2005; 28: 401–7. https://doi.org/10.1016/j.tins.2005.06.006 PMID: 15982753

198. Znamenskiy P, Zador AM. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. Nature. 2013; 497: 482–485. https://doi.org/10.1038/nature12077 PMID: 23636333

199. Friedman A, Homma D, Gibb LG, Amemori KI, Rubin SJ, Hood AS, et al. A corticostriatal path targeting striosomes controls decision-making under conflict. Cell. 2015; 161: 1320–1333. https://doi.org/10.1016/j.cell.2015.04.049 PMID: 26027737

200. Gremel CM, Chancey JH, Atwood BK, Luo G, Neve R, Ramakrishnan C, et al. Endocannabinoid Modulation of Orbitostriatal Circuits Gates Habit Formation. Neuron. 2016; 90: 1312–1324. https://doi.org/10.1016/j.neuron.2016.04.043 PMID: 27238866

201. Hooks BM, Papale AE, Paletzki RF, Feroze MW, Eastwood BS, Couey JJ, et al. Topographic precision in sensory and motor corticostriatal projections varies across cell type and cortical area. Nat Commun. 2018; 9. https://doi.org/10.1038/s41467-017-01881-x PMID: 29339724

202. Lee J, Wang W, Sabatini BL. Anatomically segregated basal ganglia pathways allow parallel behavioral modulation. Nat Neurosci. 2020. https://doi.org/10.1038/s41593-020-00712-5 PMID: 32989293