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

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

1. Department of Cognitive, Linguistic & Psychological Sciences, Brown University, USA
2. Department of Experimental Psychology, Ghent University, Belgium
3. Carney Institute for Brain Science, Brown University, USA

**Abbreviated title:** Action sequence flexibility

**Corresponding author:**
Dr. Cristian Buc Calderon
Department of Cognitive, Linguistics & Psychological Sciences
Brown University
Metcalf Research Building, 190 Thayer St, Providence, RI 02912
United States of America
Phone: +1 401-863-2727
E-mail: cbuccald@gmail.com

**# of pages:** 40
**# of figures and tables**: 7 figures, 1 table
**# of words:** abstract (209), introduction (1387), methods (2044), results (2372), discussion (2351)

**Conflict of interest:** The authors declare no competing financial interests

**Acknowledgements:** C.B.C is supported by FWO grant #12O7719N. T.V. is supported by BOF17-GOA-004. M.J.F. is supported by R01MH084840-08A1. We thank the members of the Frank and Verguts lab for helpful discussions, and Jose Miguel Buc Chavez for the bossa nova rhythm description.

**Keywords:** Action sequences, recurrent neural network, premotor cortex, basal ganglia, motor timing

## 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 order 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 order 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 and then flexibly play it in a "bossa nova" rhythm without further training.

# 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 modify the rhythm within each sequence, i.e. speed up or slow down the tempo; or apply different rhythms on a previously learned sequence, i.e. perform a rock song with a bossa nova rhythm. This implies that musicians can quickly and flexibly manipulate action timing in action sequences. Similar capabilities abound in many 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 (Jin et al., 2009; Gouvêa et al., 2015; Mello et al., 2015; Bakhurin et al., 2017; Dhawale et al., 2017), hippocampus (Pastalkova et al., 2008; MacDonald et al., 2013; Eichenbaum, 2014) and the cortex (Luczak et al., 2007; Harvey et al., 2012; Remington et al., 2018) are thought to provide a temporal (ordinal) signal for these action sequences to emerge. However, the mechanistic and dynamic principles by which these neural patterns afford sequential flexibility remain unknown. While several neural network models of corticostriatal circuit 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 (Graybiel, 1998).

In this paper, 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.

Existing 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 features of action sequences.

- In associative chain models (also termed synfire chain; e.g., Fiete et al., 2010), activation flows sequentially from one neuron (or neuronal population) to another through feedforward connections (e.g., Cone and Shouval, 2021). The sequence emerges from the hard-wired

structure of the chain. Associative chain models naturally produce sequential but also persistent neural activity, both of which are observed empirically (Veliz-Cuba et al., 2015; Pereira and Brunel, 2020). They can also deal with inherent compression of sequential activity, and thereby learn to produce each action in the sequence at any desired precise time (Cone and Shouval, 2021). 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 (Goodbody and Wolpert, 1998; Shmuelof et al., 2012). Indeed, a musician who has learned a novel rhythm can directly speed up or slow down this tempo without any additional learning. 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.

- Cluster-based models also involve a chained sequence of activation, but this sequence is learned via cell assemblies (i.e. clusters) that form within a recurrent neural network (RNN) through, for instance, spike timing dependent plasticity (Murray and Escola, 2017; Maes et al., 2020). Depending on the timing of the sequential input to the distinct subsets of the RNN, connectivity may emerge within and between clusters. Once this connectivity matrix is learned, input to the RNN induces a sequential activation whereby activation flows from one cluster to another. In contrast to associative chain models, cluster-based models allow temporal rescaling (Murray and Escola, 2017) while also producing sequential and persistent patterns of activity (Maes et al., 2020). 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) will be activated sequentially (and so forth). Therefore, the RNN can encode multiple sequential behaviors by learning (and selectively activating) distinct cluster chains encoded in the RNN connectivity matrix (Murray and Escola, 2017). 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, see Cone and Shouval, 2021). Indeed, cluster-based models can flexibly manipulate sequences; however, these sequences are typically iso-synchronous (see Murray and Escola, 2017).

- State-space models (e.g., Hardy et al., 2018) do not assume a chaining structure at all. Based on a sparsely connected RNN structure, these models are able to learn and reproduce (in the presence of a noise) a neural trajectory represented in high-dimensional space (Sussillo and Abbott, 2009; Laje and Buonomano, 2013; Rajan et al., 2016). This neural trajectory acts as a travelling wave which can then be decoded by downstream neurons to produce sequential orders. State-space models have the ability to learn highly complex and flexible motor sequences. However, unlike the other models, state-space models typically require highly supervised (i.e. continuous teaching signal) and non-biological learning mechanisms. Moreover, they do not provide a potential mechanism for encoding multiple sequences and fail at implementing temporal rescaling (unless resorting to very specific learning regimes, see Hardy et al., 2018).

- 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 order are represented jointly within the recurrent weights of the network[1]. In related sequential decision-making contexts in

[1] In associative chain models, action timing is principally controlled by the strength of recurrent connections of each excitatory neuron pool, which controls delay times for the next action in line (Cone and Shouval, 2021). Similarly, in cluster models, order and timing information (i.e., when activation "jumps" from one cluster to the other) also depends on the RNN weights (Maes et al., 2020). Finally, state-space models contain both order and time information within the same sets of RNN weights, which define both the neural trajectory and the speed at which the trajectory unfolds (Rajan et al., 2016; Remington et al., 2018).

the reinforcement learning domain, such joint coding of task features facilitates only very rigid forms of generalization and transfer, whereas the ability to code task features compositionally facilitates more robust transfer (Franklin and Frank, 2018) that can better account for human behavior (Franklin and Frank, 2020). However, the mechanisms for such compositionality in neural networks remains unknown.

Here, we develop a biologically plausible 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. To increase computational flexibility, the network factorizes order and timing information by storing them separately in a premotor cortical RNN which is dynamically gated by a basal ganglia-thalamus module. This modularity thereby affords independent (and flexible) manipulation of sequence order and action timing. For instance, once an action sequence has been learned, temporal rescaling can be accomplished by targeting the locus representing time, while still allowing the network to produce the same desired action order sequence.

In the remainder of the paper we first present the architecture of the model, and ground it within the context of neurophysiological observations on the premotor cortex (PMC) and the BG. Second, we describe how the model *learns* to produce precisely timed, temporally asynchronous, action sequences. Third, we demonstrate how the mechanistic properties of the model can account for temporal properties: temporal shifting, rescaling, compositionality, and sustained motor activation. Fourth, we simulate both empirical and neurophysiological observations. Finally, we discuss the characteristics and abilities of the ACDC model in light of neurophysiological evidence and alternative neurocomputational models.

# Methods

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

In this section, we provide the reader with an intuitive functioning of the ACDC model in simplified and more detailed forms (Fig. 1); we refer to Appendix A for a comprehensive mathematical formulation. A context module encodes the sequences to be executed (e.g., which song is to be played), and is provided as input to an RNN, which learns to encode sequence order via Hebbian learning. Order is encoded as a sequence of attractor states represented by persistent activation in distinct excitatory RNN unit clusters. Cell assemblies (or clusters) learn to project to the appropriate action identity (again via Hebbian learning), represented topographically in the BG. In turn, the BG project to the corresponding

motor thalamus to control action execution. To optimize precise action timing, the weights between action identity and execution are learned via supervised learning (i.e. Delta rule), perhaps summarizing the role of cerebellum in error corrective learning. Finally, thalamic activity about the executed action is fed back to the RNN, ultimately creating a cortico-basal ganglia loop. Each loop subtends the appropriate action order, identity and timing execution, allowing precisely timed action sequences to unfold.

In a more detailed manner, our ACDC model contains four main modules (Fig. 1 right panel): 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 figure 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 $i^{th}$ 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; Amit, 1988; Durstewitz et al., 2000). In prefrontal cortical – BG models, such clusters are referred to as “stripes” based on their anatomical existence, and are independently gated by BG (O’Reilly and Frank, 2006). 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 positions in phonological loop tasks(O’Reilly and Frank, 2006). However, in ACDC such clusters emerge naturally via learning rather than hard-coded anatomical entities. Moreover, attractor states are maintained via E-I balance: 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 E-I balance is not perturbed by another input (see below), activation in the cluster will persist and the RNN will continue representing the $i^{th}$ 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 $j^{th}$ action associated to the $i^{th}$ order (see Frank, 2006 and Ratcliff and Frank, 2012 for related computational models of evidence accumulation in these units, and Doi et al., 2020 for empirical data). Striatal Go cells, via the basal ganglia direct pathway machinery (Alexander and Crutcher, 1990; Mink, 1996), 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).

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 (Frank, 2006; Wiecki and Frank, 2013; see also Lo and Wang, 2006). 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; Pereira and Brunel, 2020).

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 (Schmitt et al., 2017; Rikhye et al., 2018), enabling the RNN to transition from the current state to the next. That is, the activation of action nodes perturbs the E-I RNN balance in a way that allows the $i^{th}$ cluster to shut down and the $i+1^{th}$ cluster to be expressed (see Appendix A for further details). 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 (Taverna et al., 2008; Wiecki and Frank, 2013; Dobbs et al., 2016), thereby shutting down evidence in favor of the $j^{th}$ action, and hence stopping the execution of the $j^{th}$ action. This loop is then reproduced with the $i^{th+1}$ RNN cluster and $j^{th+1}$ G-A-N triplet in the BG-thalamus unit, and so forth until the action sequence is performed in its entirety.

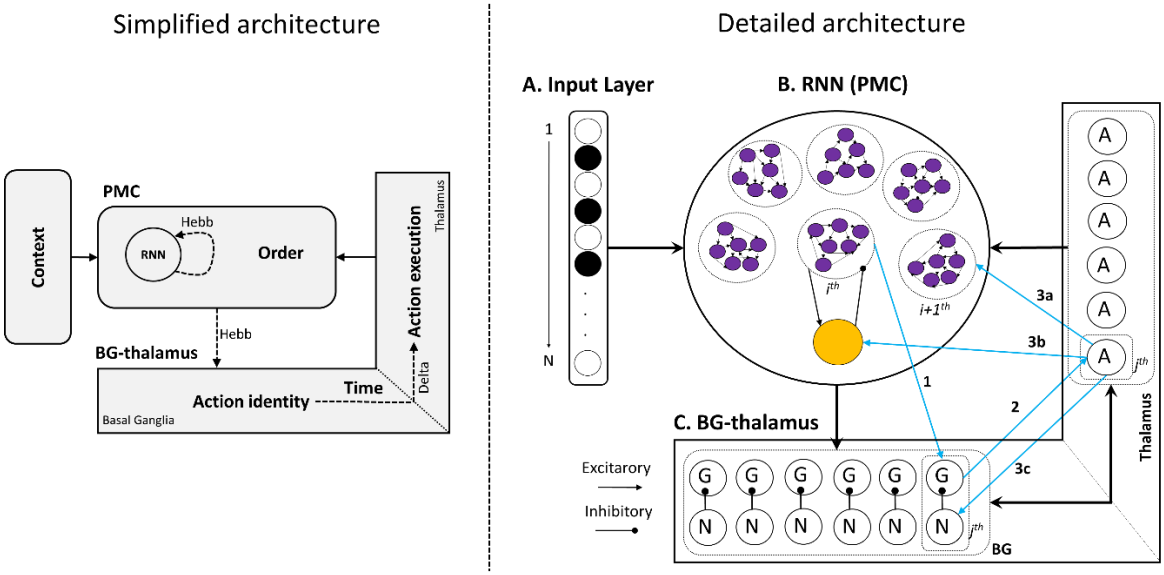



**Figure 1. Left panel: Simplified ACDC model architecture.** An input context layer indicates which sequence needs to be learned or executed. The premotor cortex (PMC) is subtended by a RNN that learns (via Hebbian learning) to form clusters of excitatory neurons encoding order in the sequence, and which are regulated by an inhibitory pool. 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. **Right panel: ACDC full model architecture. 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 the Go nodes 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 $j^{th}$ Action node selectively projects excitatory connections to the $i^{th+1}$ cluster in the RNN, the shared inhibitory neuron and the $j^{th}$ No Go node in the BG. Light blue arrows represent the $i^{th}$ cortico-basal ganglia loop instance.

Several features of the model should be highlighted. First, each cluster activation within the RNN acts as an attractor state representing the $i^{th}$ order in the sequence. Interestingly, cells in the monkey PMC code for the position in sequence, regardless of the actual movement produced during that position (Clower and Alexander, 1998; Shima and Tanji, 2000; Isoda and Tanji, 2003, 2004; Averbeck et al., 2006; Berdyyeva and Olson, 2009; Salinas, 2009). 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 (Jin et al., 2009; Schwartze et al., 2011; Gershman et al., 2014; Jones and Jahanshahi, 2014; Mello et al., 2015; Thura and Cisek, 2017; Paton and Buonomano, 2018). 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 below, these separate routes will become important for providing timing and rhythm flexibility.

Third, as in many cortico-BG models (e.g., Gurney et al., 2001; Frank, 2006), and motivated by anatomical data (Alexander et al., 1986) 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 (Oh et al., 2014; Hintiryan et al., 2016; Hunnicutt et al., 2016), whereby causal activation of specific subregions is related to specific output behaviors (Peters et al., 2021), and is also supported by human neuroimaging (Gerardin et al., 2003) and monkey/rodent neurophysiology studies (McHaffie et al., 2005; Jin et al., 2009; Znamenskiy and Zador, 2013; Friedman et al., 2015; Gremel et al., 2016; Hooks et al., 2018; Lee et al., 2020). 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 (Maes et al., 2020). 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:

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

where $\bar{x}_j$ is presynaptic activity low-pass filtered over a time scale $\tau_w$; $x_i$ is postsynaptic activity; $\alpha_1$ and $\alpha_2$ are learning rate parameters. When $\bar{x}_j$ and $x_i$ are both simultaneously > 0, $W_{ij}$ goes to *Wmax*; otherwise $W_{ij}$ goes to 0. Note that $\bar{x}_j(t)$ will be non-zero if unit *j* is active within the time window from *t* - $\tau_w \rightarrow t$ (as in Murray and Escola, 2017).

Second, Equation 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, as described in equation 2:

$$\Delta W = \eta(t_{observed} - t_{desired}) \quad \text{equation 2}$$

where the change in weight ($\Delta W$) between the $j^{th}$ Go and Action nodes is driven by the learning rate $\eta$, 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 (i.e., in chunks): the model first learns to produce the first action at the appropriate time (i.e. until the error < ϕ and ϕ is a low value, see table 1 in Appendix A), then the second, and so forth. Note that

299 learning by chunking improves motor execution (Wymbs et al., 2012; Boutin et al., 2013; Lungu et al.,
300 2014; Graybiel and Grafton, 2015; Doyon et al., 2017), and chunk-based representation is at the base of
301 several theoretical models of motor sequence learning (Abrahamse et al., 2013; Verwey et al., 2014;
302 Diedrichsen and Kornysheva, 2015).

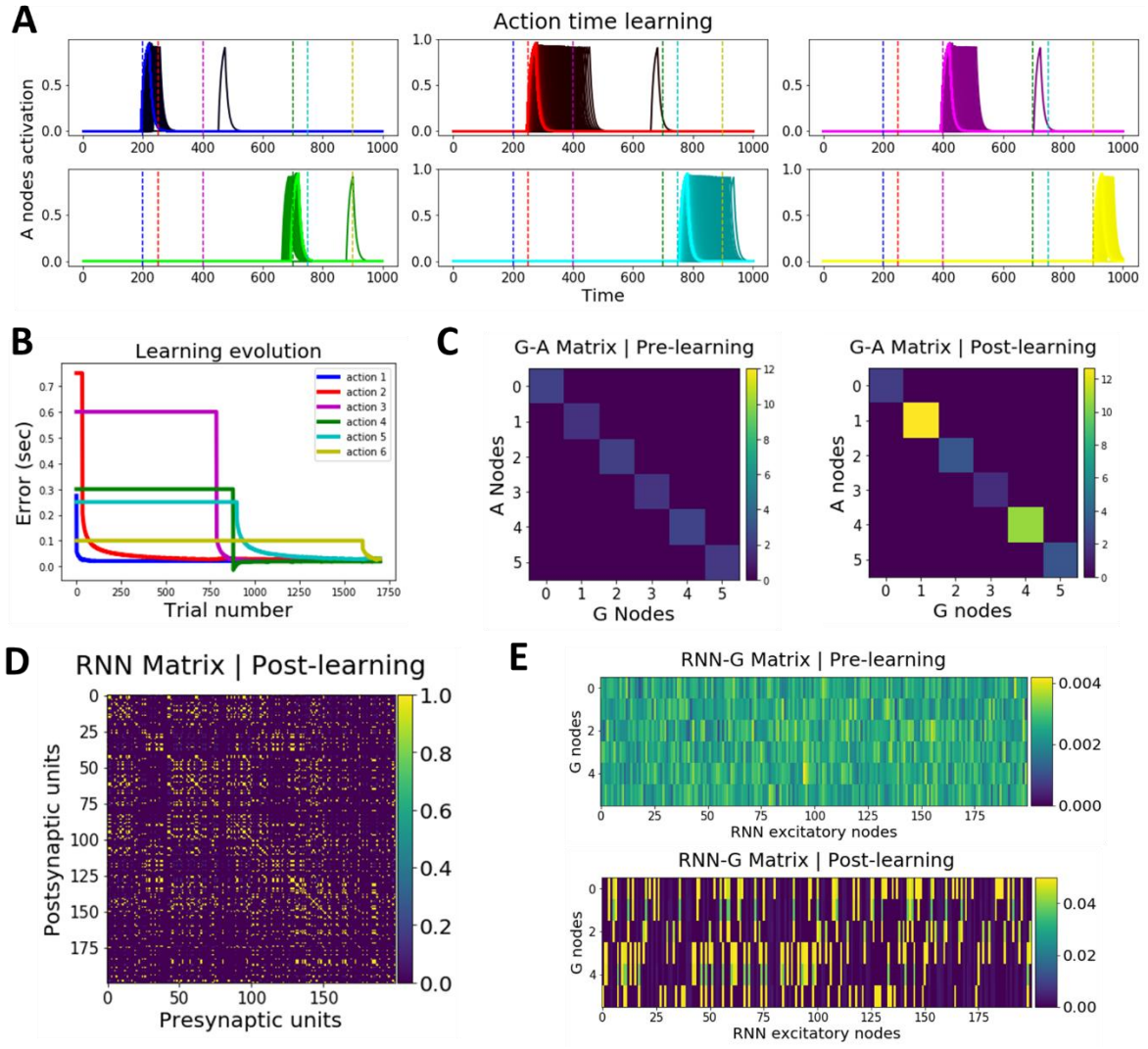
303 Below we describe all the simulations emerging from the ACDC model; parameter values for all
304 simulations are reported in table 1 of Appendix A, and simulation code is available from
305 https://github.com/CristianBucCalderon/ACDC. We start by describing how the model can learn to
306 produce precise spatiotemporal sequences. We then simulate all the temporal properties of the model:
307 reproduction of an action sequence with temporal asynchrony, temporal shifting, rescaling, and
308 compositionality, and sustained motor activation. Finally, we describe and simulate empirical and
309 neurophysiological observations.

310 # Results

311 ## *Learning precise spatiotemporal sequences*

312 Figure 2 shows the result of the first simulation, where the ACDC model learns to produce a precisely
313 timed, temporally asynchronous, action sequence. For the purpose of clarity, we limit the sequence to 6
314 actions. The goal of the model in this simulation is to produce each action sequentially at the
315 appropriate time, i.e. action 1 through 6 have to be executed respectively at times 200, 250, 400, 700,
316 750 and 900 ms (within a 1 second window). Note that this is an arbitrarily chosen timing sequence; the
317 model can (learn to) produce any timed, synchronous (see Thunderstruck simulation below) or
318 asynchronous, sequence. Figure 2A shows how the activity of each Action node goes down the gradient
319 and progressively reaches the optimal time (depicted by color coded vertical dashed lines), reflected in a
320 decrease in the action timing error (Fig. 2B) and in the weight changes between Go and Action nodes
321 (Fig. 2C).

322 Figure 2D depicts the RNN connectivity matrix after learning (weights are zero before learning).
323 Excitatory projections to the RNN from the input and motor layer are pseudo-random, with the
324 restriction that two different projections never excite the same RNN neuron. These pseudo-random
325 projections make it hard to visually identify the presence of clusters in fig2D; importantly however, this
326 connectivity matrix does induce clustered dynamics (see Video 2 and Fig. 7A below). Finally, figure 2E
327 shows how the $i^{th}$ cluster in the RNN learns to be (almost) selectively wired with the $j^{th}$ Go node.




328

329 **Figure 2. ACDC's learning dynamics. A. Learning a precisely timed action sequence.** Each action
330 execution (A node activation) is progressively shifted towards the optimal action time (depicted by the
331 color coded vertical dashed line; x-axis represents time). Learning progresses from darker to brightest
332 colors. **B. Learning evolution.** Color coded traces represent the evolution of the error as a function of
333 trial number for each action in the sequence. Learning unfolds sequentially, whereby timing errors are t
334 minimized for the first action before the second action starts learning. Therefore, each action (except
335 action 1) starts off with a plateaued error level until the preceding action reaches the optimal time.
336 Some action timings are learned faster than others because their optimal time weight value is closer to
337 their initial value. The error is computed by subtracting the observed from the desired response time
338 and plotted in seconds. **C. BG weights encode time.** Action timing is learned by changing the weights
339 from BG Go nodes to thalamus Action nodes. The left and right panel show respectively the weights
340 values before and after learning. For instance, the second action (red trace in B) starts off being
341 produced too slowly. Hence, weights increase until they produce the optimal action time for action 2.
342 Color bars indicate weight values. **D. RNN connectivity matrix after learning.** The RNN connectivity
343 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 Video 2 and Fig. 7A 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.** 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.

## *Temporal flexibility properties of the ACDC model*

Having established learned clusters within the RNN and learned sequences in the ACDC model, we now focus on the flexibility properties of the 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 now clamp learning (i.e., freeze the weights). We provide the network with the same input and observe that the network can reproduce the sequence maintaining its precision in action timing. Figure 3A shows that, given the clamped set of learned weights, each Action node (color coded for order) within the thalamus layer gets activated at the previously learnt precise timing, in a phasic/transient fashion.

**Simulation 3: temporal shifting.** The previous action sequence can be shifted in time, i.e., started 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 either adding an additional 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 (Forstmann et al., 2008; although similar effects could be implemented by

dopaminergic modulation; see Discussion). 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. Figure 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, figure 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. Naturally, intrinsic temporal constraints of the model limit this distance for earlier shifts, as shown by the negative plateau in figure 3C (black circles).

**Simulation 4: Temporal rescaling.** Musicians possess the ability to 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 ($\rho$) to all Go nodes simultaneously; if $\rho > 1$ or $0 < \rho < 1$ the sequence is respectively compressed or dilated. Figure 3D shows temporal rescaling for $\rho$ 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 $\rho$, 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. Figure 3E shows that this sum of rations stays constant (mean = 7.5, s.d. = 0.12), thereby indicating that temporal structure is maintained albeit rescaled.

**Simulation 5: Temporal compositionality.** Musicians must also be capable of temporal compositionality; that is, apply a different 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 the brain can extract and store a tempo, which then can be used as a dynamical multiplicative signal to all Go nodes as in simulation 4. In simulation 5, we apply a dynamical multiplicative signal (Fig. 3F right panel) to the Go nodes. The result is to produce the learned sequence (described in Fig. 3A) to the tempo described by the multiplicative signal. Figure 3F (left panel) shows how the time of each action in the sequence does not fall on the previous tempo (color coded vertical dashed lines), but now rather is produced at the novel timing (vertical solid lines).

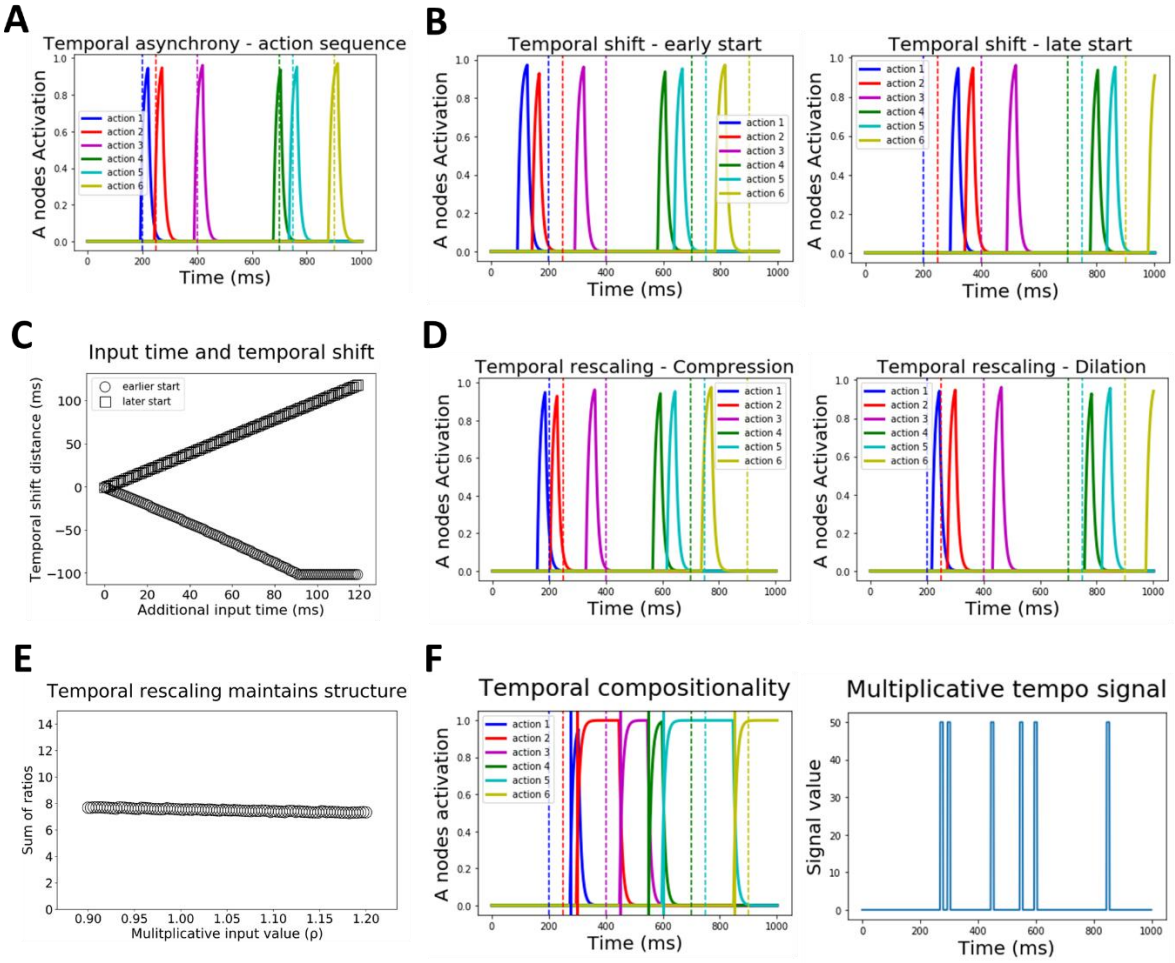



**Figure 3. Temporal properties of the ACDC model. A. Simulation 2: Reproduction of action sequence with temporal asynchrony.** Each action (i.e. A node activation, color coded) is produced at the precise desired time indicated by the vertical dashed line (also color coded), within a 1 second time window. 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 $\rho$ (x-axis). The sum of ratios value (black circles) stays constant as a function of $\rho$, 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.

**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 activation is needed. On the left panel of figure 4, we show the results of applying such a signal during the period between the start of the second action and the beginning of the third one. The motor activation of the second Action node (red trace) is sustained until the third action is executed (purple trace). Second, via a learning mechanism, the weight value between a specific Action-No Go nodes pair can be decreased to induce sustained activation of the Action node. On the right panel of figure 4, we decreased the weight value connecting the third Action-No go nodes pair. Such a weight change produced a similar result to that of implementing a multiplicative input to the No Go node, i.e. sustained activation of the corresponding Action node.

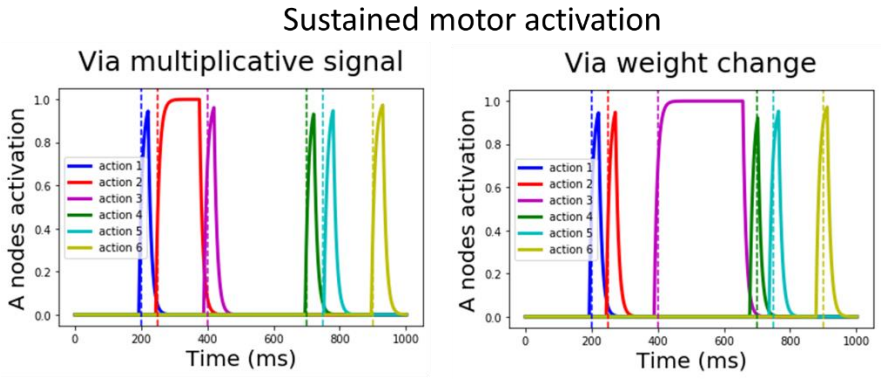



**Figure 4. Simulation 6: Sustained motor activation.** Both panels demonstrate that the ACDC model is able to output sustained motor activation as desired within a sequence. The left panel shows the results of applying a multiplicative signal ($\rho$ = 0.1) to the second No Go node, inducing a sustained 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).

**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 (Audio file 1). Notably, video 1 shows that the RNN dynamics (during the song) 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 Video 2 below). Next, we imposed the ACDC model to play the riff but now based on a bossa nova tempo without further training (Fig. 5B). The ACDC model was able to flexibly reproduce the riff following the bossa nova tempo (Fig. 5C and Audio file 2). The model thereby displays the ability to produce complex temporal compositionality. 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 Audio file 2).

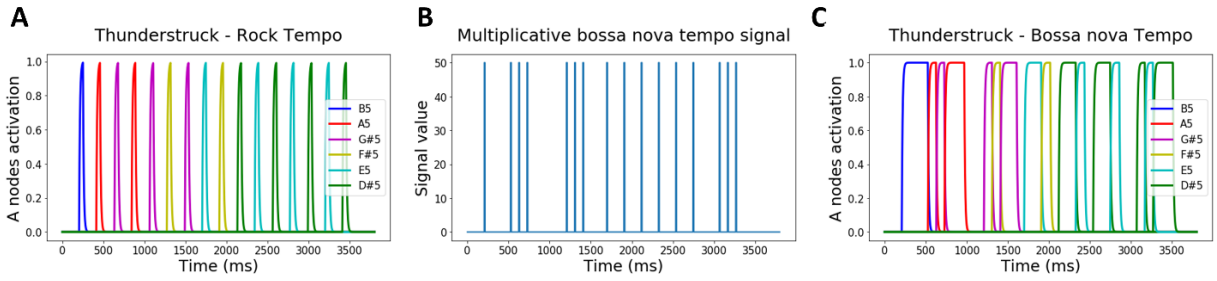



**Figure 5. Simulation 7: the ACDC produces the Thunderstruck song. A. Second guitar riff from ACDC's (the group) Thunderstruck song.** The riff is composed of 16 sequential actions creating a isosynchronous rock rhythm over a window of 3500 ms (given a 140 bpm tempo). Each action is associated to a 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.

--Insert Video 1—

**Video 1. 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 visualization of RNN dynamics as a neural trajectory in PC space. The neural 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.

## *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 (Ivry and Hazeltine, 1995; Rakitin et al., 1998; Jazayeri and Shadlen, 2010; Acerbi et al., 2012). 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 (i.e. gaussian random noise with zero mean and SD of 0.01 or 0.05 is added to the model equations 3-5 and 7). As predicted by empirical work, figure 6 shows how the SD of RTs increases as a function of interval timing for both noise values, and thereby demonstrates that the ACDC model displays scalar variability (see also Egger et al., 2020). 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, Ratcliff and Rouder, 1998). The underlying reason is that the effect of noise on evidence accumulation is amplified as time elapses.

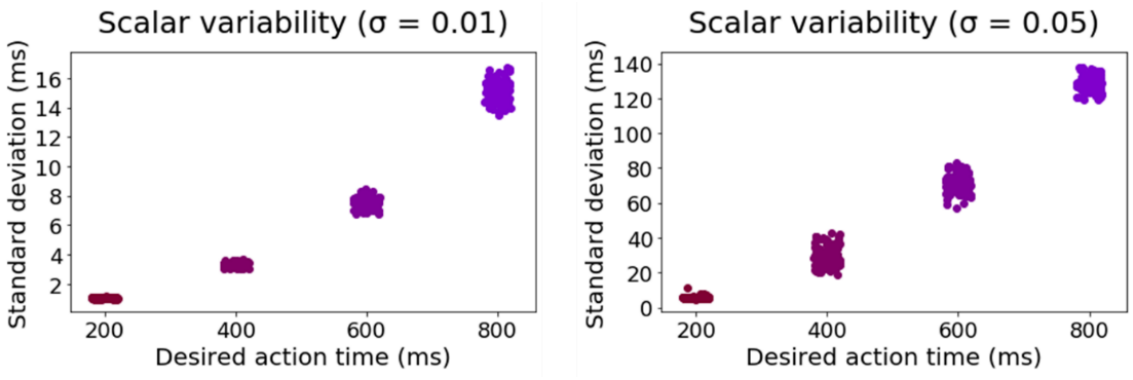



**Figure 6. Simulation 8: The ACDC model displays scalar variability.** Left (low noise value = 0.01) and right (high noise value = 0.05) panels show that the standard deviation of RTs increases as a function of the desired action time (i.e. interval timing). Moreover, higher noise values increase the range of standard deviation. Each dot is the result of 1 out of 100 simulations for each interval timing.

**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 (Miyashita and Chang, 1988; Nakamura and Kubota, 1995; Erickson and Desimone, 1999), parietal (Koch and Fuster, 1989; Chafee and Goldman-Rakic, 1998; Gail and Andersen, 2006; Klaes et al.,

2011), premotor (Cisek and Kalaska, 2005) and prefrontal (Funahashi et al., 1989, 1990; Miller et al., 1996) 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 state (Durstewitz et al., 2000; Wang, 2001; Brunel, 2003). Second, as motivated in the introduction, sequential activity has also been observed in distinct sequential behaviors such as spatial navigation (Eichenbaum, 2014) and bird song (Hahnloser et al., 2002; Kozhevnikov and Fee, 2007; Amador et al., 2013; Okubo et al., 2015).

Interestingly, recent work suggests that sequential switches in attractor states (and hence persistent neural activity), are associated to the timing of behavioral switches in action sequences (Recanatesi et al., 2020). 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 $j^{th}$ action associated to the $i^{th}$ order) is executed, this triggers a switch in attractor state in the RNN (via cortico-cortical projections from M1 to PMC), as empirically observed by Recanatesi et. al (2020). 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). Figure 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.

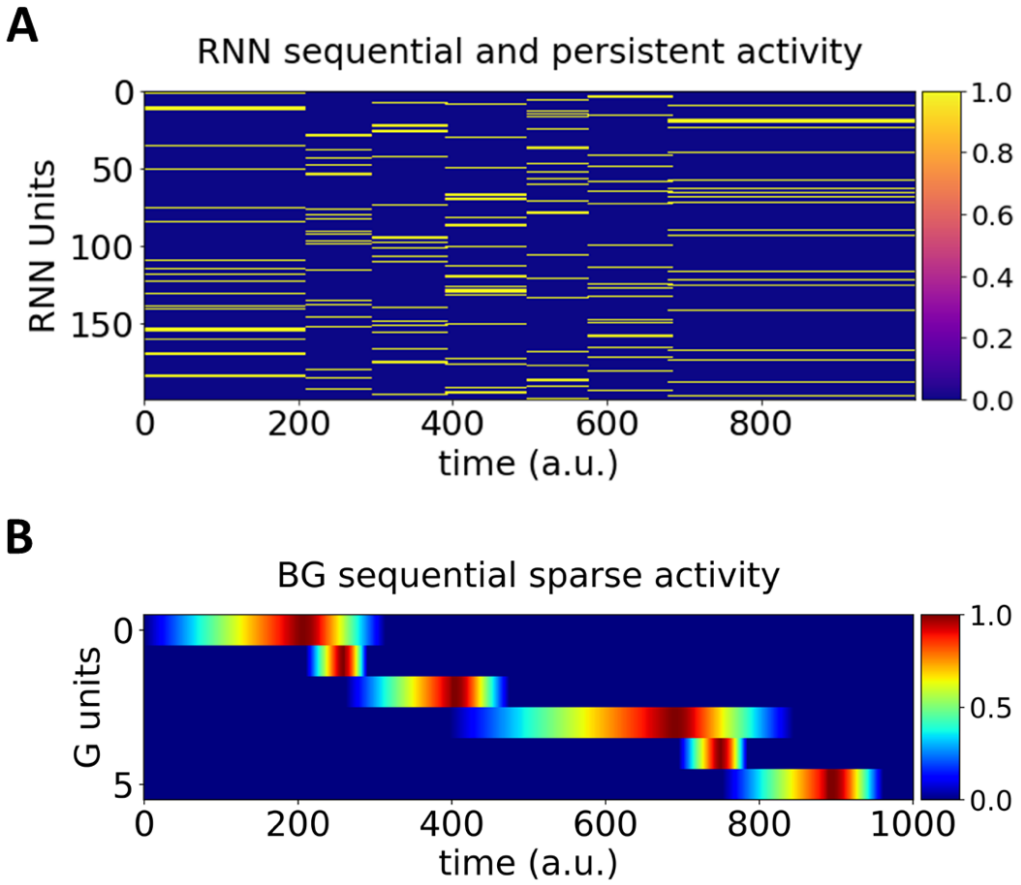



**Figure 7. Simulation 9: A. Sequential and persistent activation of clustered neural populations within the RNN.** The y-axis represents each RNN unit, the x-axis represents time. The first cluster is activated by the input layer, and maintains persistent activity until the first action is executed. At that moment, via excitatory projections from the Action nodes (Fig. 1C) to the following ($i$+1$^{th}$) 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., Gouvêa et al., 2015). Color bar represents normalized firing rate.

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 figure 7A. We then dynamically plotted the first 3 principal components (PCs) as a function of time. Video 2 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.

--Insert Video 2--

**Video 2. 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 figure 1, the right panel displays the neural RNN trajectory showing that each action execution triggers a switch from the $i^{th}$ to the $i^{th+1}$ attractor state.

The qualitative pattern of the RNN sequential and persistent dynamics (Fig. 7A) is different than observed in rodent (Harvey et al., 2012; Eichenbaum, 2014; Mello et al., 2015) or monkey (Jin et al., 2009) 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 figures 2A, 3B, 1E, 8C, respectively of Mello et al., 2015; Rueda-Orozco and Robbe, 2015; Bakhurin et al., 2017; Dhawale et al., 2017)

# Discussion

We have presented a neurocomputational model combining the strength of associative chains (e.g., Pereira and Brunel, 2020) and cluster-dependent (e.g., Maes et al., 2020) models, while also providing a model of 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 (i.e., constant or varying inter action intervals). 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.

### *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 (Ju and Bassett, 2020). Within this framework, one can envision action sequences as neural states unfolding over time. By analyzing the neural activity of secondary motor cortex in rodents, Recanaseti et al. (2020) showed that sequential behavior was subtended by the sequential unfolding of

attractor states, whereby each action in the sequence was subtended by a particular attractor state. Furthermore, these authors were able to model variability in action timing by adding correlated noise to the dynamics of a RNN. This addition allowed their RNN to display dynamics that jump from one attractor state to another, but at random times (hence explaining the variability in action timing). Our model is based on a similar mechanism. The dynamical activity of the RNN reflects switches in attractor states induced by excitatory projections to the RNN that transiently modify its E-I balance. However, via BG learning and modulation, our model precisely controls the switch to another cortical attractor state, thereby influencing output timing. Within our conceptualization, we suggest that persistent activity within a cluster indicates the latent state that the system is in (e.g., Collins and Frank, 2013), which in this case reflects the ordinal position in the sequence. Moreover, in contrast to previous models, the 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. The nature of this activity can take the form of a cyclical signal (Hartley and Houghton, 1996), decaying start signal (Page and Norris, 1998), or a sequence of overlapping states (Burgess and Hitch, 1999, 2006). Other models assume that the network input (used to learn the sequence) is itself sequential in nature (Murray and Escola, 2017; Maes et al., 2020), 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 (Kawai et al., 2015), and transiently inactivated or disturbed via optogenetic manipulation (Otchy et al., 2015). 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 (Koziol and Budding, 2009; Collins and Frank, 2016a).

### *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 (Remington et al., 2018). Murray and Escola (2017) 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 Egger et al., 2020; Kozachkov and Michmizos, 2020)(see also Egger et al., 2020). However, a recent model making use of eligibility traces (Florian, 2007; Izhikevich, 2007; Frémaux et al., 2010; Soltoggio and Steil, 2013; Bellec et al., 2019), manages to learn precise non-isosynchronous spatiotemporal sequence learning (Cone and Shouval, 2021). Still, it is unclear how such a model can rescale non-isosynchronous sequences, and neither of these models is capable of exhibiting temporal compositionality. Crucially, the ACDC model can perform temporal rescaling for both isocynchronous and non-isosynchronous sequencing, and it can also flexibly switch the tempo altogether through a multiplicative signal to the BG.

The temporal properties of our model discussed in the previous paragraph 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 (Watabe-Uchida et al., 2017), and several studies have implicated dopamine in controlling movement vigor (Beierholm et al., 2013; Hamid et al., 2015, 2021; Panigrahi et al., 2015; Zénon et al., 2016; Berke, 2018; Gaidica et al., 2018; Sedaghat-Nejad et al., 2019; Augustin et al., 2020). Dopamine has also been extensively implicated in impulsive (i.e. pathologically speeded) behavior (van Gaalen et al., 2006; Frank et al., 2007; Pattij and Vanderschuren, 2008; Buckholtz et al., 2010; Pine et al., 2010; Dalley and Roiser, 2012; Economidou et al., 2012). 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 (Lake and Meck, 2013).

If dopamine can flexibly modulate (i.e. 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 (Ratcliff, 1978; Ratcliff and Rouder, 1998), 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 (Yousif et al., 2016; Beste et al., 2018) or reward-based (Westbrook et al., 2020) decision-making tasks. Our model implements this possibility. Indeed, inputs to Go nodes modify (i.e. 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 (Heitz, 2014). Therefore, a second alternative is that dopamine or other BG modulations may modify the threshold of action execution (Wiecki and Frank, 2013; Lloyd and Dayan, 2015). Interestingly, Parkinson's disease patients on subthalamic deep brain stimulation tend to behave impulsively (Frank et al., 2007), due to modulation of the decision threshold (Frank, 2006; Cavanagh et al., 2011; Herz et al., 2016). Naturally, both hypotheses are not mutually exclusive; further research should investigate the effects of dopaminergic and subthalamic modulations regarding motor sequence flexibility.

### *Biological basis and learning*

In line with recent models (Murray and Escola, 2017; Maes et al., 2020; Cone and Shouval, 2021), the ACDC model implements a certain level of biological plausibility, and still is able to capture a plethora of data both at the neurophysiological and behavioral level. For instance, we demonstrate that the model can replicate sequential sparse activation observed within the basal ganglia (Gouvêa et al., 2015). Another model making use of RNNs set at a near chaotic regime (Rajan et al., 2016) has been able to replicate sparse sequential activations as recorded in mice parietal cortex during spatial navigation (Harvey et al., 2012). Yet, training in these networks is based on highly supervised mechanisms that are not biologically plausible (Sussillo and Abbott, 2009; Laje and Buonomano, 2013; Hardy et al., 2018). Therefore, future research should analyze whether more biological plausible RNNs (see Miconi, 2017) can reproduce such action patterns. Note that some of the implementation details of our model have still to be worked out (see limitations section below).

### *Encoding and executing multiple sequences*

One important advantage of cluster-based models is the potential to encode multiple sequences within the same network of interconnected neurons. Within our model, this would tantamount to having several sequential attractor state neural trajectories, each of which subtends the execution of one specific action sequence. Therefore, action sequence selection is seen as targeting a specific cluster of units within the RNN, leading to the execution of the corresponding sequence. Moreover, the ability to produce various sequences simultaneously would resume to simply activating more than one cluster (Murray and Escola, 2017).

Another important question focuses on investigating how sequences are chained one after the other in order to produce adaptive behavior. Neurophysiological recordings in mice have revealed the existence of specific neural codes during sequences that signal the beginning and end (or boundaries) of a sequence (Jin et al., 2014; Jin and Costa, 2015). These go and stop signal may be used to signal the system to transition from one neural trajectory to the other, thereby allowing action sequences to be chained (Logiaco et al., 2019).

### *Limitations and future directions*

As previously noted, some of the implementation details of our modem have still to be worked out. For instance, reinforcement learning of action timing is conceptually thought to take the form of a tri-factor hebbian learning rule (Montague et al., 1995; Bailey et al., 2000; Izhikevich, 2007; Hoerzer et al., 2014; Miconi, 2017), where neurons subtending a rewarding behavior (and hence forming a specific cortical activity patterns) increase their connectivity to D1 receptor striatal populations (also known as Go cells) via dopaminergic activity bursts stemming from midbrain nuclei (Frank, 2005; Collins and Frank, 2013). Our implementation is slightly different. Reinforcement is shifted later in the information processing pipeline, and action time learning takes place between Go nodes (which we also consider as D1 receptor cells in the striatum) and thalamic motor neurons. We applied a delta rule within the BG-thalamus module. Much evidence indicates that the BG learn via reinforcement learning (e.g., McClure et al., 2003; O'Doherty et al., 2004; Badre and Frank, 2012), but the brain also makes use of signed errors for precise timing e.g., in the cerebellum. Our learning rule in the BG-thalamus thus summarizes the contributions of these systems in conjunction. In contrast, classical Hebbian learning rule was applied within the RNN and between the RNN and BG. Indeed, these projections simply carry “chaining” information (i.e. they allow for the sequential structure of the chain to emerge during learning), and therefore do not need to be fine-tuned to a specific value for the emergence of precise action timing. Future work should consider both biological learning constraints and implement a more detailed architecture of the basal ganglia networks.

Moreover, future work on the ACDC should focus on investigating the limits with which RNNs can encode multiple sequences without creating interference. This entails exploring two aspects of the model. First, our model forces orthogonalization of the inputs to the RNN in order to make sure that no clusters are interconnected. Novel versions of the ACDC need to investigate how this orthogonalization may emerge from specific learning rules. Second, cluster size mater, as bigger cluster size may be more robust to noise. Therefore, given an initial number of RNN excitatory neurons, only but a limited amount

of sequences can be encoded. Therefore, the interaction between learning, cluster size and sequence interference should also be investigated more systematically.

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 (Erlhagen and Schöner, 2002; Cisek, 2006; Klaes et al., 2012), which have been shown to successfully model more continuous reaching actions (Christopoulos et al., 2015). Moreover, these continuous action representations in the BG may require additional inputs from the cerebellum for movement coordination (Thach et al., 1992) or sequence prediction for motor control (Bastian, 2006).

Finally, recent research focused on how humans extract abstract knowledge, and generalize this knowledge to other situations (Collins and Frank, 2016b) (Collins and Frank, 2013; Whittington et al., 2020; Baram et al., 2021). 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 (Lehnert et al., 2020).

# Conclusion

Separating time and order information in two distinct loci of a biologically inspired model of action sequences allowed us to increase computational power, and to capture a significant amount of data at the neurophysiological and behavioral levels. Although some specific aspect of the implementation of this functional specialization need still to be resolved, we demonstrate that such an architecture increased motor flexibility. We propose a concrete and tractable mechanism of this flexibility, and thus suggest a model of how humans and animals can learn and effortlessly manipulate precise spatiotemporal signals at the basis of complex behavior.

# References

Abrahamse EL, Ruitenberg MFL, de Kleine E, Verwey WB (2013) Control of automated behavior: Insights from the discrete sequence production task. Front Hum Neurosci 7:1–16.

Acerbi L, Wolpert DM, Vijayakumar S (2012) Internal Representations of Temporal Statistics and Feedback Calibrate Motor-Sensory Interval Timing. PLoS Comput Biol 8.

Alexander GE, Crutcher MD (1990) Functional architecture of basal ganglia circuits: neural substrates of parallel processing. Trends Neurosci 13:266–271.

Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci VOL. 9:357–381.

Amador A, Perl YS, Mindlin GB, Margoliash D (2013) Elemental gesture dynamics are encoded by song premotor cortical neurons. Nature 495:59–64.

Amit DJ (1988) Neural networks counting chimes. Proc Natl Acad Sci U S A 85:2141–2145.

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

Averbeck BB, Sohn JW, Lee D (2006) Activity in prefrontal cortex during dynamic selection of action sequences. Nat Neurosci 9:276–282.

Badre D, Frank MJ (2012) Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: evidence from fMRI. Cereb Cortex 22:527–536 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3278316&tool=pmcentrez&rendertype=abstract [Accessed December 12, 2014].

Bailey CH, Giustetto M, Huang YY, Hawkins RD, Kandel ER (2000) Is Heterosynaptic modulation essential for stabilizing hebbian plasiticity and memory. Nat Rev Neurosci 1:11–20.

Bakhurin KI, Goudar V, Shobe JL, Claar LD, Buonomano D V., Masmanidis SC (2017) Differential Encoding of Time by Prefrontal and Striatal Network Dynamics. J Neurosci 37:854–870 Available at: http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.1789-16.2017.

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

Bastian AJ (2006) Learning to predict the future: the cerebellum adapts feedforward movement control. Curr Opin Neurobiol 16:645–649.

Beierholm U, Guitart-Masip M, Economides M, Chowdhury R, Düzel E, Dolan R, Dayan P (2013) Dopamine modulates reward-related vigor. Neuropsychopharmacology 38:1495–1503.

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

Berdyyeva TK, Olson CR (2009) Monkey Supplementary Eye Field Neurons Signal the Ordinal Position of Both Actions and Objects. J Neurosci 29:591–599 Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.4803-08.2009.

Berke JD (2018) What does dopamine mean? Nat Neurosci 21:787–793 Available at: http://dx.doi.org/10.1038/s41593-018-0152-y.

Beste C, Adelhöfer N, Gohil K, Passow S, Roessner V, Li SC (2018) Dopamine modulates the efficiency of sensory evidence accumulation during perceptual decision making. Int J Neuropsychopharmacol 21:649–655.

Boutin A, Massen C, Heuer H (2013) Modality-specific organization in the representation of sensorimotor sequences. Front Psychol 4:1–9.

Brunel N (2003) Dynamics and Plasticity of Stimulus-selective Persistent Activity in Cortical Network Models. Cereb Cortex 13:1151–1161.

Buckholtz JW, Treadway MT, Cowan RL, Woodward ND, Li R, Ansari MS, Baldwin RM, Schwartzman AN, Shelby ES, Smith CE, Kessler RM, Zald DH (2010) Dopaminergic network differences in human impulsivity. Science (80- ) 329:532.

Burgess N, Hitch GJ (1999) Memory for serial order: A network model of the phonological loop and its timing. Psychol Rev 106:551–581 Available at: http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-295X.106.3.551.

Burgess N, Hitch GJ (2006) A revised model of short-term memory and long-term learning of verbal sequences. J Mem Lang 55:627–652.

Cavanagh JF, Wiecki T V, Cohen MX, Figueroa CM, Samanta J, Sherman SJ, Frank MJ (2011) Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. Nat Neurosci 14:1462–1467 Available at: http://www.ncbi.nlm.nih.gov/pubmed/21946325.

Chafee M V., Goldman-Rakic PS (1998) Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. J Neurophysiol 79:2919–2940.

Christopoulos V, Bonaiuto J, Andersen R a. (2015) A Biologically Plausible Computational Theory for Value Integration and Action Selection in Decisions with Competing Alternatives. PLOS Comput Biol 11:e1004104 Available at: http://dx.plos.org/10.1371/journal.pcbi.1004104.

Cisek P (2006) Integrated neural processes for defining potential actions and deciding between them: a computational model. J Neurosci 26:9761–9770 Available at: http://www.ncbi.nlm.nih.gov/pubmed/16988047 [Accessed August 9, 2013].

Cisek P, Kalaska JF (2005) Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. Neuron 45:801–814 Available at: http://www.ncbi.nlm.nih.gov/pubmed/15748854 [Accessed March 2, 2013].

Clower WT, Alexander GE (1998) Movement sequence-related activity reflecting numerical order of components in supplementary and presupplementary motor areas. J Neurophysiol 80:1562–1566.

Collins AGE, Frank MJ (2013) Cognitive control over learning: creating, clustering, and generalizing task-set structure. Psychol

789 Rev 120:190–229 Available at: http://doi.apa.org/getdoi.cfm?doi=10.1037/a0030852.

790 Collins AGE, Frank MJ (2016a) Motor Demands Constrain Cognitive Rule Structures. PLoS Comput Biol 12:1–17.

791 Collins AGE, Frank MJ (2016b) Neural signature of hierarchically structured expectations predicts clustering and transfer of rule
792 sets in reinforcement learning. Cognition 152:160–169 Available at: http://dx.doi.org/10.1016/j.cognition.2016.04.002.

793 Cone I, Shouval HZ (2021) Learning precise spatiotemporal sequences via biophysically realistic learning rules in a modular,
794 spiking network. Elife 10:2020.04.17.046862 Available at: https://elifesciences.org/articles/63751.

795 Dalley JW, Roiser JP (2012) Dopamine, serotonin and impulsivity. Neuroscience 215:42–58 Available at:
796 http://dx.doi.org/10.1016/j.neuroscience.2012.03.065.

797 Dhawale AK, Poddar R, Wolff SBE, Normand VA, Kopelowitz E, Ölveczky BP (2017) Automated long-Term recording and analysis
798 of neural activity in behaving animals. Elife 6:1–40.

799 Diedrichsen J, Kornysheva K (2015) Motor skill learning between selection and execution. Trends Cogn Sci 19:227–233.

800 Dobbs LKK, Kaplan ARR, Lemos JCC, Matsui A, Rubinstein M, Alvarez VAA (2016) Dopamine Regulation of Lateral Inhibition
801 between Striatal Neurons Gates the Stimulant Actions of Cocaine. Neuron 90:1100–1113 Available at:
802 http://dx.doi.org/10.1016/j.neuron.2016.04.031.

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

805 Doi T, Fan Y, Gold JI, Ding L (2020) The caudate nucleus contributes causally to decisions that balance reward and uncertain
806 visual information. Elife 9:568733 Available at: https://www.biorxiv.org/content/10.1101/568733v1.

807 Doyon J, Gabitov E, Vahdat S, Lungu O, Boutin A (2017) ScienceDirect Current issues related to motor sequence learning in
808 humans. Curr Opin Behav Sci 20:89–97 Available at: http://dx.doi.org/10.1016/j.cobeha.2017.11.012.

809 Durstewitz D, Seamans JK, Sejnowski TJ (2000) Neurocomputational Models of Working Memory. Nat Neurosci 3:1184–1191.

810 Economidou D, Theobald DEH, Robbins TW, Everitt BJ, Dalley JW (2012) Norepinephrine and dopamine modulate impulsivity on
811 the five-choice serial reaction time task through opponent actions in the shell and core sub-regions of the nucleus
812 accumbens. Neuropsychopharmacology 37:2057–2066.

813 Egger SW, Le NM, Jazayeri M (2020) A neural circuit model for human sensorimotor timing. Nat Commun 11:1–14 Available at:
814 http://dx.doi.org/10.1038/s41467-020-16999-8.

815 Eichenbaum H (2014) Time cells in the hippocampus: A new dimension for mapping memories. Nat Rev Neurosci 15:732–744
816 Available at: http://dx.doi.org/10.1038/nrn3827.

817 Erickson CA, Desimone R (1999) Responses of macaque perirhinal neurons during and after visual stimulus association learning.
818 J Neurosci 19:10404–10416.

819 Erlhagen W, Schöner G (2002) Dynamic field theory of movement preparation. Psychol Rev 109:545–572 Available at:

http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-295X.109.3.545 [Accessed December 3, 2014].

Fiete IR, Senn W, Wang CZH, Hahnloser RHR (2010) Spike-Time-Dependent Plasticity and Heterosynaptic Competition Organize Networks to Produce Long Scale-Free Sequences of Neural Activity. Neuron 65:563–576 Available at: http://dx.doi.org/10.1016/j.neuron.2010.02.003.

Florian R V. (2007) Reinforcement learning through modulation of spike-timing-dependent synaptic plasticity. Neural Comput 19:1468–1502.

Forstmann BU, Dutilh G, Brown SD, Neumann J, von Cramon DY, Ridderinkhof KR, Wagenmakers E-J (2008) Striatum and pre-SMA facilitate decision-making under time pressure. Proc Natl Acad Sci U S A 105:17538–17542.

Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J Cogn Neurosci 17:51–72.

Frank MJ (2006) Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. Neural Netw 19:1120–1136 Available at: http://www.sciencedirect.com/science/article/pii/S089360800600150X.

Frank MJ, Samanta J, Moustafa AA, Sherman SJ (2007) Hold your horses: impulsivity, deep brain stimulation, and medication in parkinsonism. Science 318:1309–1312 Available at: http://www.ncbi.nlm.nih.gov/pubmed/17962524.

Franklin NT, Frank MJ (2018) Compositional clustering in task structure learning Daunizeau J, ed. PLOS Comput Biol 14:e1006116 Available at: https://dx.plos.org/10.1371/journal.pcbi.1006116.

Franklin NT, Frank MJ (2020) Generalizing to generalize: Humans flexibly switch between compositional and conjunctive structures during reinforcement learning. Available at: http://dx.doi.org/10.1371/journal.pcbi.1007720.

Frémaux N, Sprekeler H, Gerstner W (2010) Functional requirements for reward-modulated spike-timing-dependent plasticity. J Neurosci 30:13326–13337.

Friedman A, Homma D, Gibb LG, Amemori KI, Rubin SJ, Hood AS, Riad MH, Graybiel AM (2015) A corticostriatal path targeting striosomes controls decision-making under conflict. Cell 161:1320–1333 Available at: http://dx.doi.org/10.1016/j.cell.2015.04.049.

Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol 61:331–349.

Funahashi S, Bruce CJ, Goldman-Rakic PS (1990) Visuospatial coding in primate prefrontal neurons revealed by oculomotor paradigms. J Neurophysiol 63:814–831.

Gaidica M, Hurst A, Cyr C, Leventhal DK (2018) Distinct populations of motor thalamic neurons encode action initiation, action selection, and movement vigor. J Neurosci 38:6563–6573.

Gail A, Andersen R a (2006) Neural dynamics in monkey parietal reach region reflect context-specific sensorimotor transformations. J Neurosci 26:9376–9384 Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1570-06.2006%5Cnpapers3://publication/doi/10.1523/JNEUROSCI.1570-06.2006.

852 Gerardin E, Lehéricy S, Pochon JB, Du Montcel ST, Mangin JF, Poupon F, Agid Y, Le Bihan D, Marsault C (2003) Foot, hand, face
853 and eye representation in the human striatum. Cereb Cortex 13:162–169.

854 Gershman SJ, Moustafa AA, Ludvig EA (2014) Time representation in reinforcement learning models of the basal ganglia. Front
855 Comput Neurosci 7:1–8.

856 Goodbody SJ, Wolpert DM (1998) Temporal and amplitude generalization in motor learning. J Neurophysiol 79:1825–1838.

857 Gouvêa TS, Monteiro T, Motiwala A, Soares S, Machens C, Paton JJ (2015) Striatal dynamics explain duration judgments. Elife
858 4:1–14.

859 Graybiel AM (1998) The Basal Ganglia and Chunking of Action Repertoires. Neurobiol Learn Mem 70:119–136 Available at:
860 https://linkinghub.elsevier.com/retrieve/pii/S1074742798938436.

861 Graybiel AM, Grafton ST (2015) The striatum: Where skills and habits meet. Cold Spring Harb Perspect Biol 7:1–14.

862 Gremel CM, Chancey JH, Atwood BK, Luo G, Neve R, Ramakrishnan C, Deisseroth K, Lovinger DM, Costa RM (2016)
863 Endocannabinoid Modulation of Orbitostriatal Circuits Gates Habit Formation. Neuron 90:1312–1324 Available at:
864 http://dx.doi.org/10.1016/j.neuron.2016.04.043.

865 Gurney K, Prescott TJ, Redgrave P (2001) A computational model of action selection in the basal ganglia. I. A new functional
866 anatomy. Biol Cybern 84:401–410.

867 Hahnloser RHR, Kozhevnikov AA, Fee MS (2002) An ultra-sparse code underlies the generation of neural sequences in a
868 songbird. Nature 419:65–70.

869 Hamid AA, Frank MJ, Moore CI (2021) Dopamine waves as a mechanism for spatiotemporal credit assignment. Cell.

870 Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, Berke JD (2015)
871 Mesolimbic dopamine signals the value of work. Nat Neurosci 19:117–126.

872 Hardy NF, Goudar V, Romero-Sosa JL, Buonomano D V. (2018) A model of temporal scaling correctly predicts that motor timing
873 improves with speed. Nat Commun 9:1–14 Available at: http://dx.doi.org/10.1038/s41467-018-07161-6.

874 Hartley T, Houghton G (1996) A linguistically constrained model of short-term memory for nonwords. J Mem Lang 35:1–31.

875 Harvey CD, Coen P, Tank DW (2012) Choice-specific sequences in parietal cortex during a virtual-navigation decision task.
876 Nature 484:62–68.

877 Heitz RP (2014) The speed-accuracy tradeoff: History, physiology, methodology, and behavior. Front Neurosci 8:1–19.

878 Herz DM, Zavala BA, Bogacz R, Brown P (2016) Neural Correlates of Decision Thresholds in the Human Subthalamic Nucleus.
879 Curr Biol 26:916–920 Available at: http://linkinghub.elsevier.com/retrieve/pii/S0960982216001330.

880 Hintiryan H, Foster NN, Bowman I, Bay M, Song MY, Gou L, Yamashita S, Bienkowski MS, Zingg B, Zhu M, Yang XW, Shih JC, Toga
881 AW, Dong HW (2016) The mouse cortico-striatal projectome. Nat Neurosci 19:1100–1114.

882 Hoerzer GM, Legenstein R, Maass W (2014) Emergence of complex computational structures from chaotic neural networks

through reward-modulated hebbian learning. Cereb Cortex 24:677–690.

Hooks BM, Papale AE, Paletzki RF, Feroze MW, Eastwood BS, Couey JJ, Winnubst J, Chandrashekar J, Gerfen CR (2018) Topographic precision in sensory and motor corticostriatal projections varies across cell type and cortical area. Nat Commun 9 Available at: http://dx.doi.org/10.1038/s41467-018-05780-7.

Hunnicutt BJ, Jongbloets BC, Birdsong WT, Gertz KJ, Zhong H, Mao T (2016) A comprehensive excitatory input map of the striatum reveals novel functional organization. Elife 5:1–32.

Isoda M, Tanji J (2003) Contrasting Neuronal Activity in the Supplementary and Frontal Eye Fields during Temporal Organization of Multiple Saccades. J Neurophysiol 90:3054–3065.

Isoda M, Tanji J (2004) Participation of the primate presupplementary motor area in sequencing multiple saccades. J Neurophysiol 92:653–659.

Ivry RB, Hazeltine RE (1995) Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. J Exp Psychol Hum Percept Perform 21:3–18 Available at: http://doi.apa.org/getdoi.cfm?doi=10.1037/0096-1523.21.1.3.

Izhikevich EM (2007) Solving the distal reward problem through linkage of STDP and dopamine signaling. Cereb Cortex 17:2443–2452.

Jazayeri M, Shadlen MN (2010) Temporal context calibrates interval timing. Nat Neurosci 13:1020–1026.

Jin DZ, Fujii N, Graybiel AM (2009) Neural representation of time in cortico-basal ganglia circuits. Proc Natl Acad Sci U S A 106:19156–19161.

Jin X, Costa RM (2015) Shaping action sequences in basal ganglia circuits. Curr Opin Neurobiol 33:188–196 Available at: http://dx.doi.org/10.1016/j.conb.2015.06.011.

Jin X, Tecuapetla F, Costa RM (2014) Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. Nat Neurosci 17:423–430.

Jones CRG, Jahanshahi M (2014) Contributions of the Basal Ganglia to Temporal Processing: Evidence from Parkinson's Disease. Timing Time Percept 2:87–127.

Ju H, Bassett DS (2020) Dynamic representations in networked neural systems. Nat Neurosci 23:908–917 Available at: http://dx.doi.org/10.1038/s41593-020-0653-3.

Kawai R, Markman T, Poddar R, Ko R, Fantana AL, Dhawale AK, Kampff AR, Ölveczky BP (2015) Motor Cortex Is Required for Learning but Not for Executing a Motor Skill. Neuron 86:800–812.

Klaes C, Schneegans S, Schöner G, Gail A (2012) Sensorimotor Learning Biases Choice Behavior: A Learning Neural Field Model for Decision Making. PLoS Comput Biol 8:e1002774 Available at: http://dx.doi.org/10.1371/journal.pcbi.1002774%5Cnhttp://www.ploscompbiol.org/article/fetchObjectAttachment.action?uri=info:doi/10.1371/journal.pcbi.1002774&representation=PDF.

915 Klaes C, Westendorff S, Chakrabarti S, Gail A (2011) Choosing goals, not rules: deciding among rule-based action plans. Neuron
916 70:536–548 Available at: http://www.ncbi.nlm.nih.gov/pubmed/21555078 [Accessed March 7, 2013].

917 Koch KW, Fuster JM (1989) Unit activity in monkey parietal cortex related to haptic perception and temporary memory. Exp
918 Brain Res.

919 Kozachkov L, Michmizos KP (2020) Sequence Learning in Associative Neuronal-Astrocytic Networks. In: Lecture Notes in
920 Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp
921 349–360 Available at: http://link.springer.com/10.1007/978-3-030-59277-6_32.

922 Kozhevnikov AA, Fee MS (2007) Singing-related activity of identified HVC neurons in the zebra finch. J Neurophysiol 97:4271–
923 4283.

924 Koziol LF, Budding DE (2009) Subcortical structures and cognition: Implications for neuropsychological assessment. Springer
925 Science & Business Media.

926 Laje R, Buonomano D V (2013) Robust timing and motor patterns by taming chaos in recurrent neural networks. Nat Neurosci
927 16:925–933 Available at: http://www.nature.com/doifinder/10.1038/nn.3405.

928 Lake JI, Meck WH (2013) Differential effects of amphetamine and haloperidol on temporal reproduction: Dopaminergic
929 regulation of attention and clock speed. Neuropsychologia 51:284–292 Available at:
930 http://dx.doi.org/10.1016/j.neuropsychologia.2012.09.014.

931 Lee J, Wang W, Sabatini BL (2020) Anatomically segregated basal ganglia pathways allow parallel behavioral modulation. Nat
932 Neurosci Available at: http://dx.doi.org/10.1038/s41593-020-00712-5.

933 Lehnert L, Littman ML, Frank MJ (2020) Reward-predictive representations generalize across tasks in reinforcement learning.
934 PLoS Comput Biol 16:1–27 Available at: http://dx.doi.org/10.1371/journal.pcbi.1008317.

935 Lloyd K, Dayan P (2015) Tamping Ramping: Algorithmic, Implementational, and Computational Explanations of Phasic Dopamine
936 Signals in the Accumbens. PLoS Comput Biol 11:1–34.

937 Lo CC, Wang XJ (2006) Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. Nat Neurosci
938 9:956–963.

939 Logiaco L, Abbott LF, Escola S (2019) A model of flexible motor sequencing through thalamic control of cortical dynamics.
940 bioRxiv:2019.12.17.880153.

941 Luczak A, Barthó P, Marguet SL, Buzsáki G, Harris KD (2007) Sequential structure of neocortical spontaneous activity in vivo.
942 Proc Natl Acad Sci U S A 104:347–352.

943 Lungu O, Monchi O, Albouy G, Jubault T, Ballarin E, Burnod Y, Doyon J (2014) Striatal and hippocampal involvement in motor
944 sequence chunking depends on the learning strategy. PLoS One 9:25–27.

945 MacDonald CJ, Carrow S, Place R, Eichenbaum H (2013) Distinct hippocampal time cell sequences represent odor memories in
946 immobilized rats. J Neurosci 33:14607–14616.

947 Maes A, Barahona M, Clopath C (2020) Learning spatiotemporal signals using a recurrent spiking network that discretizes time.
948 PLoS Comput Biol 16:1–26 Available at: http://dx.doi.org/10.1371/journal.pcbi.1007606.

949 Mcclure SM, Berns GS, Montague PR (2003) Temporal Prediction Errors in a Passive Learning Task Activate Human Striatum.
950 38:339–346.

951 McHaffie JG, Stanford TR, Stein BE, Coizet V, Redgrave P (2005) Subcortical loops through the basal ganglia. Trends Neurosci
952 28:401–407 Available at: http://www.ncbi.nlm.nih.gov/pubmed/15982753 [Accessed November 21, 2014].

953 Mello GBM, Soares S, Paton JJ (2015) A scalable population code for time in the striatum. Curr Biol 25:1113–1122 Available at:
954 http://dx.doi.org/10.1016/j.cub.2015.02.036.

955 Miconi T (2017) Biologically plausible learning in recurrent neural networks reproduces neural dynamics observed during
956 cognitive tasks. Elife 6:1–24.

957 Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. J
958 Neurosci 16:5154–5167.

959 Mink JW (1996) The basal ganglia: Focused selection and inhibition of competing motor programs. Prog Neurobiol 50:381–425.

960 Miyashita Y, Chang HS (1988) Neuronal correlate of pictorial short-term memory in the primate temporal cortex. Nature
961 331:68–70 Available at: http://www.nature.com/articles/331068a0.

962 Montague PR, Dayan P, Person C, Sejnowski TJ (1995) Bee foraging in uncertain environments. Nature 377:725–728 Available
963 at: http://papers.cnl.salk.edu/PDFs/Bee Foraging in Uncertain Environments Using Predictive Hebbian Learning 1995-
964 3013.pdf.

965 Murray JM, Escola GS (2017) Learning multiple variable-speed sequences in striatum via cortical tutoring. Elife 6:1–24.

966 Nakamura K, Kubota K (1995) Mnemonic firing of neurons in the monkey temporal pole during a visual recognition memory
967 task. J Neurophysiol 74:162–178.

968 O'Reilly RC, Frank MJ (2006) Making working memory work: a computational model of learning in the prefrontal cortex and
969 basal ganglia. Neural Comput 18:283–328 Available at: http://cognet.mit.edu/journal/10.1162/089976606775093909.

970 Oh SW et al. (2014) A mesoscale connectome of the mouse brain. Nature 508:207–214.

971 Okubo TS, Mackevicius EL, Payne HL, Lynch GF, Fee MS (2015) Growth and splitting of neural sequences in songbird vocal
972 development. Nature 528:352–357.

973 Otchy TM, Wolff SBE, Rhee JY, Pehlevan C, Kawai R, Kempf A, Gobes SMH, Ölveczky BP (2015) Acute off-target effects of neural
974 circuit manipulations. Nature 528:358–363.

975 Page MPA, Norris D (1998) The primacy model: A new model of immediate serial recall. Psychol Rev 105:761–781 Available at:
976 https://psycnet.apa.org/buy/1998-12733-005.

977 Panigrahi B, Martin KA, Li Y, Graves AR, Vollmer A, Olson L, Mensh BD, Karpova AY, Dudman JT (2015) Dopamine Is Required for

the Neural Representation and Control of Movement Vigor. Cell 162:1418–1430.

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

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

Pattij T, Vanderschuren LJMJ (2008) The neuropharmacology of impulsive behaviour. Trends Pharmacol Sci 29:192–199.

Pereira U, Brunel N (2020) Unsupervised Learning of Persistent and Sequential Activity. Front Comput Neurosci 13:1–19.

Peters AJ, Fabre JMJ, Steinmetz NA, Harris KD, Carandini M (2021) Striatal activity topographically reflects cortical activity. Nature 591 Available at: http://dx.doi.org/10.1038/s41586-020-03166-8.

Pine A, Shiner T, Seymour B, Dolan RJ (2010) Dopamine, time, and impulsivity in humans. J Neurosci 30:8888–8896.

Rajan K, Harvey CDD, Tank DWW (2016) Recurrent Network Models of Sequence Generation and Memory. Neuron 90:128–142 Available at: http://dx.doi.org/10.1016/j.neuron.2016.02.009.

Rakitin BC, Penney TB, Gibbon J, Malapani C, Hinton SC, Meck WH (1998) Scalar expectancy theory and peak-interval timing in humans. J Exp Psychol Anim Behav Process 24:15–33.

Ratcliff R (1978) A theory of memory retrieval. Psychol Rev 85:59–108.

Ratcliff R, Frank MJ (2012) Reinforcement-based decision making in corticostriatal circuits: Mutual constraints by neurocomputational and diffusion models. Neural Comput 24:1186–1229.

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

Recanatesi S, Pereira U, Murakami M, Mainen Z, Mazzucato L (2020) Metastable attractors explain the variable timing of stable behavioral action sequences. bioRxiv:0–40.

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

Rikhye R V., Gilra A, Halassa MM (2018) Thalamic regulation of switching between cortical representations enables cognitive flexibility. Nat Neurosci 21:1753–1763 Available at: http://dx.doi.org/10.1038/s41593-018-0269-z.

Rueda-Orozco PE, Robbe D (2015) The striatum multiplexes contextual and kinematic information to constrain motor habits execution. Nat Neurosci 18:453–462.

Salinas E (2009) Rank-Order-Selective Neurons Form a Temporal Basis Set for the Generation of Motor Sequences. J Neurosci 29:4369–4380 Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.0164-09.2009.

Schmitt LI, Wimmer RD, Nakajima M, Happ M, Mofakham S, Halassa MM (2017) Thalamic amplification of cortical connectivity sustains attentional control. Nature 545:219–223.

1009 Schwartze M, Keller PE, Patel AD, Kotz SA (2011) The impact of basal ganglia lesions on sensorimotor synchronization,
1010 spontaneous motor tempo, and the detection of tempo changes. Behav Brain Res 216:685–691 Available at:
1011 http://dx.doi.org/10.1016/j.bbr.2010.09.015.

1012 Sedaghat-Nejad E, Herzfeld DJ, Shadmehr R (2019) Reward prediction error modulates saccade vigor. J Neurosci 39:5010–5017.

1013 Shima K, Tanji J (2000) Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of
1014 multiple movements. J Neurophysiol 84:2148–2160 Available at: http://www.ncbi.nlm.nih.gov/pubmed/11024102.

1015 Shmuelof L, Krakauer JW, Mazzoni P (2012) How is a motor skill learned? Change and invariance at the levels of task success
1016 and trajectory control. J Neurophysiol 108:578–594.

1017 Soltoggio A, Steil JJ (2013) Solving the distal reward problem with rare correlations. Neural Comput 25:940–978.

1018 Sussillo D, Abbott LF (2009) Generating Coherent Patterns of Activity from Chaotic Neural Networks. Neuron 63:544–557
1019 Available at: http://dx.doi.org/10.1016/j.neuron.2009.07.018.

1020 Taverna S, Ilijic E, Surmeier DJ (2008) Recurrent collateral connections of striatal medium spiny neurons are disrupted in models
1021 of Parkinson's disease. J Neurosci 28:5504–5512.

1022 Thach WT, Goodkin HP, Keating JG (1992) The cerebellum and the adaptive coordination of movement. Annu Rev Neurosci
1023 15:403–442.

1024 Thura D, Cisek P (2017) The Basal Ganglia Do Not Select Reach Targets but Control the Urgency of Commitment. Neuron
1025 95:1160-1170.e5 Available at: http://dx.doi.org/10.1016/j.neuron.2017.07.039.

1026 van Gaalen MM, van Koten R, Schoffelmeer ANM, Vanderschuren LJMJ (2006) Critical Involvement of Dopaminergic
1027 Neurotransmission in Impulsive Decision Making. Biol Psychiatry 60:66–73.

1028 Veliz-Cuba A, Shouval HZ, Josić K, Kilpatrick ZP (2015) Networks that learn the precise timing of event sequences. J Comput
1029 Neurosci 39:235–254.

1030 Verwey WB, Shea CH, Wright DL (2014) A cognitive framework for explaining serial processing and sequence execution
1031 strategies. Psychon Bull Rev 22:54–77.

1032 Wang XJ (2001) Synaptic reverberation underlying mnemonic persistent activity. Trends Neurosci 24:455–463.

1033 Watabe-Uchida M, Eshel N, Uchida N (2017) Neural Circuitry of Reward Prediction Error. Annu Rev Neurosci 40:373–394.

1034 Westbrook A, van den Bosch R, Määttä JI, Hofmans L, Papadopetraki D, Cools R, Frank MJ (2020) Dopamine promotes cognitive
1035 effort by biasing the benefits versus costs of cognitive work. Science (80- ) 367:1362–1366.

1036 Whittington JCR, Muller TH, Mark S, Chen G, Barry C, Burgess N, Behrens TEJ (2020) The Tolman-Eichenbaum Machine: Unifying
1037 Space and Relational Memory through Generalization in the Hippocampal Formation. Cell 183:1249-1263.e23 Available
1038 at: https://doi.org/10.1016/j.cell.2020.10.024.

1039 Wiecki T V, Frank MJ (2013) A computational model of inhibitory control in frontal cortex and basal ganglia. Psychol Rev

120:329–355.

Wymbs NF, Bassett DS, Mucha PJ, Porter MA, Grafton ST (2012) Article Differential Recruitment of the Sensorimotor Putamen and Frontoparietal Cortex during Motor Chunking in Humans. Neuron 74:936–946 Available at: http://dx.doi.org/10.1016/j.neuron.2012.03.038.

Yousif N, Fu RZ, Abou-El-Ela Bourquin B, Bhrugubanda V, Schultz SR, Seemungal BM (2016) Dopamine activation preserves visual motion perception despite noise interference of human v5/mt. J Neurosci 36:9303–9312.

Zénon A, Devesse S, Olivier E (2016) Dopamine manipulation affects response vigor independently of opportunity cost. J Neurosci 36:9516–9525.

Znamenskiy P, Zador AM (2013) Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. Nature 497:482–485 Available at: http://dx.doi.org/10.1038/nature12077.

# Appendix A

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

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 figure 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 equation 3:

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

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. $\gamma_E$ is the gain on Action nodes activation projected to the excitatory RNN neurons (see below for the functional property of this parameter). $\Theta$, the non-linear transformation function, is governed by $\Theta(x) = (2 / (1 + e^{-\lambda x})) - 1$

(where $\lambda$ is the gain parameter and with additional non-linearity at zero, i.e. $\Theta(x) = 0$ *if* $\Theta(x) < 0$); and $\tau_{rnn}$ 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 equation 4:

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

where $J^{IE}$, $J^{IA}$ and $\gamma_I$ 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 $j^{th}$ action associated to the $i^{th}$ order, following equation 5:

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

where $g_j$ is the activation of the $j^{th}$ 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 $j^{th}$ No Go and Go nodes, $n_j$ is the activation of the $j^{th}$ No Go node, and $\tau_g$ is the encoding constant (with $\tau_g$ >>> 0, thereby simulating evidence accumulation-like dynamics).

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 equation 6:

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

where $a_j$ is the activation of the $j^{th}$ action, $g_j$ is the activation of the $j^{th}$ Go unit, $b$ is the negative bias (i.e. threshold), $\Theta$ is a nonlinear function as in equation 3, and $\tau_a$ is the encoding constant. $J^{AG}$ is the weight from the $j^{th}$ Go unit to the $j^{th}$ Action unit, and was 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 E-I RNN balance in a way that allows the $i^{th}$ 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 equation 7:

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

where $n_j$ is the activation of the $j^{th}$ No Go node, $J^{NA}$ is the weight from the $j^{th}$ Action unit to the $j^{th}$ No Go unit, $a_j$ is the activation of the $j^{th}$ Action node, and $\tau_n$ is the encoding constant.

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

Table 1. Parameter values for all simulations

| Parameters | Values |
|---|---|
| $\alpha_{1\,(RNN)}$ / $\alpha_{1\,(RNN\text{-}Go)}$ | 0.01 / 0.00002 |
| $\alpha_{2\,(RNN)}$ / $\alpha_{2\,(RNN\text{-}Go)}$ | 0.1 / 0.4 |
| $W_{max\,(RNN)}$ / $W_{max\,(RNN\text{-}Go)}$ | 1 / 0.05 |
| $\tau_w$, b | 0.5 |
| η | 0.4 |
| ϕ | 0.01 |
| $\tau_a$, $\tau_n$, $J^{IE}$ | 0.1 |
| $\gamma_E$ / $\gamma_I$ | 21.4 / 21 |
| $\tau_{rnn}$, $x^{in}$, $J^{EI}$, $J^{EA}$, $J^{IA}$, $J^{GN}$, $J^{NA}$ | 1 |
| $\lambda_{rnn}$ / $\lambda_a$ | 10 / 10000 |
| $\tau_g$ | 0.001 |