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: 35

# of figures and tables: 8 figures, 1 table

# of words: abstract (208), introduction (1303), results (3523), discussion (2371), methods (2630)

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, Laureline Logiaco and the anonymous reviewers for valuable comments, 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 sequence information (i.e., the position of the action) and timing (i.e., the moment of response execution) are typically stored in the same neural network weights. Here, we augment a biologically plausible recurrent neural network of cortical dynamics to include a basal ganglia-thalamic module which uses reinforcement learning to dynamically modulate action. This "associative cluster-dependent chain" (ACDC) model modularly stores sequence and timing information in distinct loci of the network. This feature increases computational power and allows ACDC to display a wide range of temporal properties (e.g., multiple sequences, temporal shifting, rescaling, and compositionality), while still accounting for several behavioral and neurophysiological empirical observations. Finally, we apply this ACDC network to show how it can learn the famous "Thunderstruck" song intro and then flexibly play it in a "bossa nova" rhythm without further training.

Introduction

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

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

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

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

- In <u>associative chain models</u> (also termed synfire chain[19–21]), activation flows sequentially from one neuron (or neuronal population) to another through feedforward connections[22]. The sequence emerges from the hard-wired structure of the chain. Associative chain models produce sequential and persistent neural activity, as observed empirically[23,24]. They can deal with inherent compression of sequential activity, and learn to produce each action in the

sequence at any desired precise time[22]. However, these models are not equipped to facilitate *temporal rescaling*: the finding that learned action sequences can be sped up (compressed) or slowed down (dilated) without the need to overwrite previous learning[25,26]. Moreover, it is unclear how these models implement *temporal shifting*: the ability to start the action sequence earlier or later in time, without modifying the action sequence structure. Chain models also do not straightforwardly allow networks to encode more than a single sequence, given their hard-wired nature.

- Cluster-based models also involve a chained sequence of activation, but this sequence is learned via cell assemblies (clusters) that are chained within a recurrent neural network (RNN), for instance, through spike timing dependent plasticity[27–32]. Hence, the chaining emerges from synaptic learning rather than being hard-wired. Once the chain is learnt, a single initial input pulse to the RNN induces a sequential activation whereby activation flows from one cluster to another. Cluster-based models allow temporal rescaling[28] while also producing sequential and persistent patterns of activity[27]. Furthermore, they provide a simple mechanism allowing a network to encode multiple sequences. By selectively activating a specific cluster within the RNN, only the cluster "in line" (i.e., connected to the previous cluster) is activated sequentially. Therefore, the RNN can encode multiple sequential actions by learning (and selectively activating) distinct cluster chains encoded in the connectivity matrix[28]. Yet, it is unclear how these models could facilitate action sequences with *temporal asynchrony*: the ability to learn, and flexibly manipulate, motor sequences with varying inter-action intervals (an advantage of associative chain models[22]). Indeed, cluster-based models can flexibly manipulate sequences; however, these sequences are typically iso-synchronous[28]. In addition, although emergent connectivity within and between clusters (or units) can arise via unsupervised learning[33], this connectivity crucially depends on the sequential nature and timing of the teaching input signal to the distinct subsets of the RNN.

- State-space models[34] do not assume a chaining structure at all. Based on a (sparse) randomly connected RNN structure, these models produce a neural trajectory that evolves in high-dimensional space which can be used as a temporal basis to perform a range of complex tasks[35]. However, to reliably reproduce the same task, neural trajectories need to be robust to noise. To that end, state-space models typically harness a noiseless neural trajectory (based on any

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

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

In sum, all models can account for distinct functionalities in sequence production, but fail to provide a plausible neurocomputational mechanism from which most fundamental abilities – temporal asynchrony, shifting, rescaling, compositionality – can emerge and interact. These limitations arise from a property common to all action sequence models: action identity, timing and sequence order are represented jointly within the recurrent weights of the network. In the reinforcement learning domain, such joint coding of task features facilitates only rigid forms of generalization and transfer, whereas the ability to code task features compositionally facilitates more robust transfer[47] that can better account for human behavior[48]. However, the mechanisms for such compositionality in neural networks remains unknown.

Here, we develop a biologically inspired[1] RNN called the associative cluster-dependent chain (ACDC) model. By combining strengths of the associative chain and cluster-based models, ACDC accounts for biological data. As we show below, the novelty in our model is twofold. First, we propose a biologically-plausible model of cortico-basal ganglia-thalamic loops that decomposes the functions of cortex and basal ganglia and learns sequences based on simple local and (biologically motivated) supervised learning rules. Second, this decomposition affords greater flexibility in generating desired action sequences, supporting temporal asynchrony, shifting, rescaling, and compositionality in a single model. Crucially, our model factorizes action sequence features within the circuit, with cortical RNN representing latent states within a sequence, and BG controlling both the timing of the transitions from one state to the next and which actions are linked to sequence positions. Factorizing order and timing information by storing them separately in a premotor cortical RNN, which is dynamically gated by a basal ganglia-thalamus module, affords independent (and flexible) manipulation of sequence order and action timing, and thus increases computational flexibility.

Results

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

[1] We further discuss biological plausibility of our model in the discussion.

171 and to a shared inhibitory neuron (blue arrow 3B in Fig. 1), consistent with evidence that thalamic units
172 target both cortical excitatory and inhibitory neurons[50–52]. Thus, whenever an action is executed, the
173 ratio of excitatory to inhibitory inputs to the RNN is perturbed in a way that induces a transition from
174 the current cluster to the next cluster in line (targeted by the feedback projections of the current A
175 node, blue arrow 3A in Fig. 1) to be expressed (see Methods for more details).

176 Learning takes place over fast and slow time scales. Hebbian learning is fast and unfolds within the
177 dynamics of a trial (i.e., during the evolution of an action sequence). In contrast, the delta rule is slow
178 and is implemented between trials, via a signed error computed through the discrepancy between the
179 action timing provided by the tutor and the generated action. Action sequences are learned
180 sequentially: the model learns to produce the first action at the appropriate time, then the second, and
181 so forth. Sequential learning improves motor execution[53–57], and is at the base of several theoretical
182 models of motor sequence learning[58–60].

183 At a higher-level, order is encoded as a sequence of attractor states represented by persistent activation
184 in distinct excitatory RNN unit clusters (cell assemblies). These clusters do not represent the actions
185 themselves but rather their abstract order; the specific actions to be executed are learned via RNN
186 projections to the BG and their timing is encoded in the weights of topographic projections to the motor
187 thalamus. To optimize precise action timing, the weights between action identity (G unit activity) and
188 execution (timing for a given action conditional on G unit activity) are learned via supervised learning
189 (i.e., delta rule), perhaps summarizing the role of cerebellum in error corrective learning. This allows us
190 to model tasks in which a tutor provides feedback (e.g., [61]; see methods). Finally, feedback to the RNN
191 from thalamic activity ultimately creates a cortico-basal ganglia loop. Each loop subtends the
192 appropriate action order, identity and timing execution, allowing precisely timed action sequences to
193 unfold. As we show below, our model architecture, allowing to uniquely encode timing information in a
194 distinct subset of the network (BG) than the one encoding order (PMC), will prove to display
195 advantageous properties. In particular, being able to flexibly control, via external stimulation, the
196 dynamics of the BG will result in a model displaying several temporal flexibility properties.

197

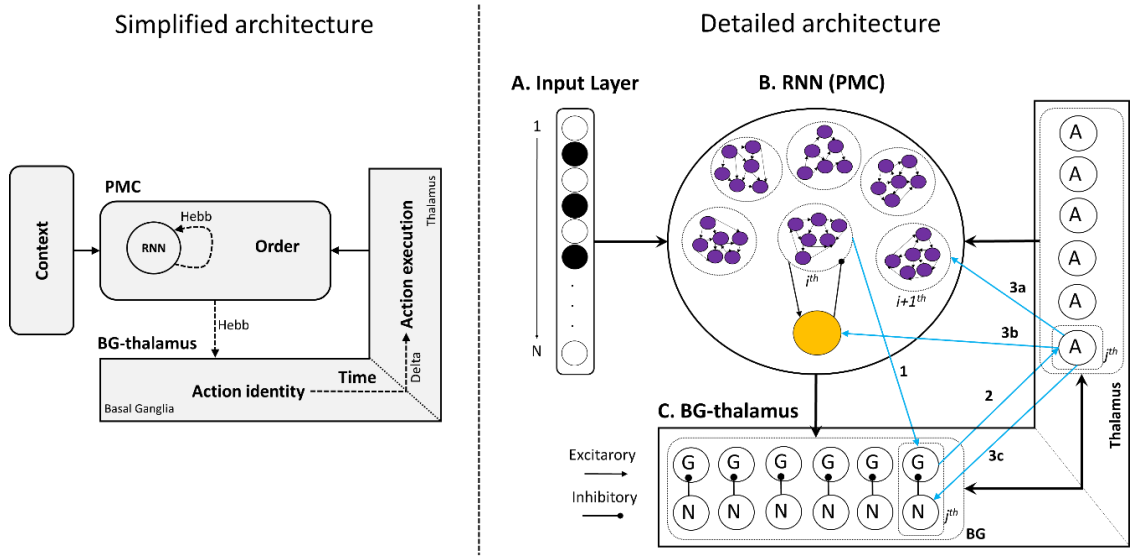

Figure 1: Simplified ACDC model architecture (left). An input context layer indicates which sequence needs to be learned or executed. The premotor cortex (PMC) is subtended by a RNN that learns (via Hebbian learning) to form clusters of excitatory neurons encoding order in the sequence, and which are regulated by an inhibitory 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. Dashed lines are plastic connections, and the associated learning rule is indicated (Hebb = hebbian learning, Delta= delta rule). Note that the plastic connections from action identity to action execution correspond solely to the blue projection n°2 on the detailed architecture figure. **ACDC full model architecture (right). A. Input layer:** codes for contexts indicating the sequence to be learned/produced in a *N* length binary vector. **B. RNN:** represents recurrently interconnected neurons of the PMC, composed of a subset of interconnected neurons (i.e., clusters) that can give rise to sequential activation states after learning via cortico-basal ganglia loops. All excitatory nodes in the RNN project to a shared inhibitory neuron (orange node), which in turn inhibits all excitatory neurons (purple nodes; shown for just one cluster for visual simplicity). **C. The BG:** composed of two neuron types G (Go cells) and N (No Go cells). Go nodes accumulate evidence over time and excite Action (A) nodes in the BG output /thalamus layer. Once activity in a Go node reaches a specific threshold, the corresponding action is executed. Once executed, Action nodes reciprocally activate No Go nodes which in turn suppress Go nodes, shutting down action execution. **The thalamus:** is composed of Action nodes whose activity represents action execution. The jth Action node selectively projects excitatory connections to the i+1th cluster in the RNN, the shared inhibitory neuron and the jth No Go node in the BG. Light blue arrows represent the ith cortico-basal ganglia loop instance. The subindex *i* refers to the ordinal position in the sequence, i.e. the order the action possesses in the sequence. The subindex *j* represents the action that is associated to the ordinal position in the sequence.

Learning precise spatiotemporal sequences

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

Figure 2D depicts the RNN connectivity matrix after learning (weights are zero before learning). Excitatory projections to the RNN from the input and motor layer are pseudo-random, with the restriction that two different projections never excite the same RNN neuron. These pseudo-random projections make it hard to visually identify the presence of clusters in figure 2D; importantly however, this connectivity matrix does induce clustered dynamics (see Video 2 and Fig. 7A below). Figure 2E shows how the ith cluster in the RNN learns to be (almost) selectively wired with the jth Go node. Finally, for completeness and transparency, figure 2F portrays the dynamics of G (left) and N (right) nodes after learning of the sequence.

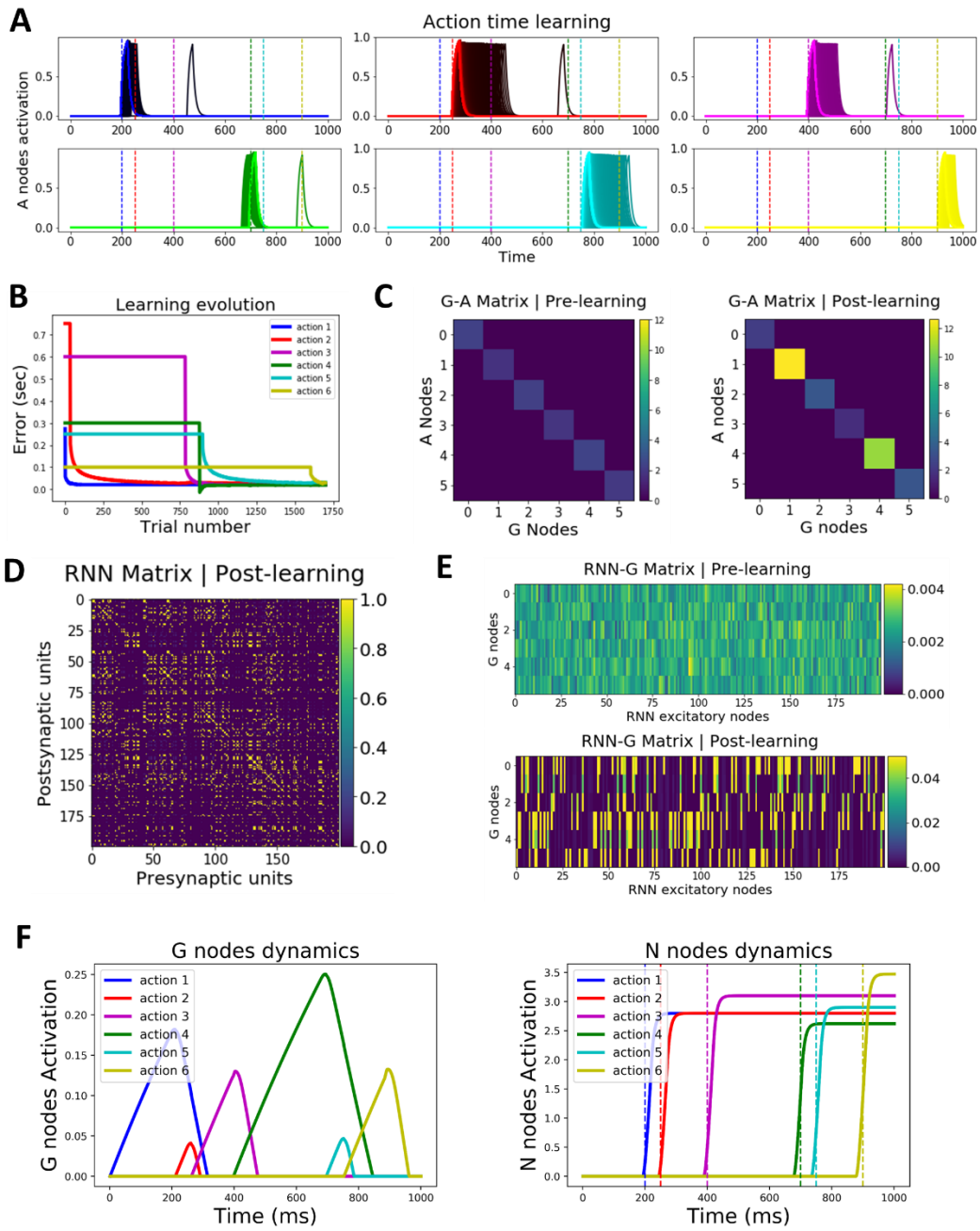

Figure 2. ACDC's learning dynamics. A. Learning a precisely timed action sequence. Each action execution (A node activation) is progressively shifted towards the optimal action time (depicted by the color coded vertical dashed line; x-axis represents time). Learning progresses from darker to brightest colors. **B. Learning evolution.** Color coded traces represent the evolution of the error as a function of trial number for each action in the sequence. Learning unfolds sequentially, whereby timing errors are t minimized for the first action before the second action starts learning. Therefore, each action (except action 1) starts off with a plateaued error level until the preceding action reaches the optimal time. Some action timings are learned faster than others because their optimal time weight value is closer to their initial value. The error is computed by subtracting the observed from the desired response time

and plotted in seconds. **C. BG weights encode time.** Action timing is learned by changing the weights from BG Go nodes to thalamus Action nodes. The left and right panel show respectively the weights values before and after learning. For instance, the second action (red trace in B) starts off being produced too slowly. Hence, weights increase until they produce the optimal action time for action 2. Color bars indicate weight values. **D. RNN connectivity matrix after learning.** The RNN connectivity matrix is initialized as a blank slate (all values are set to 0). After learning, the RNN connectivity matrix displays the appearance of clusters, whereby groups of 20 neurons are fully interconnected with each other and not connected with other neurons in the RNN (please refer to 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. **F.** Dynamics of G (left panel) and N (right panel) nodes after learning.

Temporal flexibility properties of the ACDC model

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

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

Simulation 3: temporal shifting. The previous action sequence can be shifted in time, i.e., initiated earlier or later. Importantly, this shift can occur without changing the timing between actions (i.e., sequence timing is preserved). The ACDC model achieves flexible temporal shifting by implementing an additive positive (to start the sequence earlier) or negative (to start it later) input to the first Go node of

286 the sequence, analogous to the top-down input from pre-SMA to striatum thought to bias starting
287 points for evidence accumulation[62] (although similar effects could be implemented by dopaminergic
288 modulation; see Discussion). In simulation 3, we inject an additional input of +1 or -1 to the first Go node
289 during the first 100 ms of the 1 second time window. Figure 3B shows how the sequence is shifted
290 earlier in time for the positive input (left panel) and later in time for the negative input (right panel).
291 Moreover, figure 3C shows that as this additional input lasts longer, the distance (in time) between the
292 first action of the shifted sequence and that of the original sequence increases linearly.

293 **Simulation 4: Temporal rescaling.** Musicians can learn a rhythm, i.e., a precisely timed action sequence,
294 and instantly temporally rescale (compress or dilate) that rhythm without additional learning. In our
295 model, flexible rescaling is achieved by sending a multiplicative input (ρ) to all Go nodes simultaneously
296 (i.e. the ρ parameter multiplies the net input in equation 5, see methods); if $\rho > 1$ or $0 < \rho < 1$ the
297 sequence is respectively compressed or dilated. Figure 3D shows temporal rescaling for ρ values of 1.2
298 (compression, left panel in Fig. 3D) and 0.9 (dilation, right panel in Fig. 3D). Importantly, temporally
299 rescaling the sequence does not affect the temporal structure of action sequences. For 100 values of ρ,
300 ranging from 0.9 to 1.2, we computed the relative ratio between a sequence of 3 actions. The ratio was
301 computed by subtracting the time of action 1 from that of action 2 (subtraction 1), then the time of
302 action 2 from that of action 3 (subtraction 2), and dividing subtraction 2 / subtraction 1. We performed
303 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
304 shows that this sum of rations stays constant (mean = 7.5, s.d. = 0.12), thereby indicating that temporal
305 structure is maintained albeit rescaled. Note that rescaling also induces a tiny shift in the sequence. This
306 is an emergent property of a global rescaling signal to all Go nodes of the network; this slight shift could
307 be avoided by targeting all but the first Go node with the multiplicative term.

308 **Simulation 5: Temporal compositionality.** Musicians must also be capable of temporal compositionality;
309 that is, applying a desired tempo to an action sequence that was learned in a different tempo (e.g.,
310 apply a bossa nova tempo to a rock song; see below). In simulation 5, we assume that desired tempos
311 are learned and extracted from other sequences, which then can then be used as a dynamical
312 multiplicative input signal to all Go nodes (Fig. 3F right panel). Therefore, whereas temporal rescaling
313 makes use of a constant ρ multiplicative input value to the Go nodes, temporal compositionality is
314 achieved by a dynamic multiplicative input (the ρ multiplicative input value follows the blue trace in the
315 right panel of Fig. 3F). The result is to produce the learned sequence (described in Fig. 3A) to the tempo
316 described by the multiplicative signal. Figure 3F (left panel) shows how the time of each action in the

sequence does not accord with the learned tempo (color coded vertical dashed lines), but is rather produced at the novel desired timing (vertical solid lines).

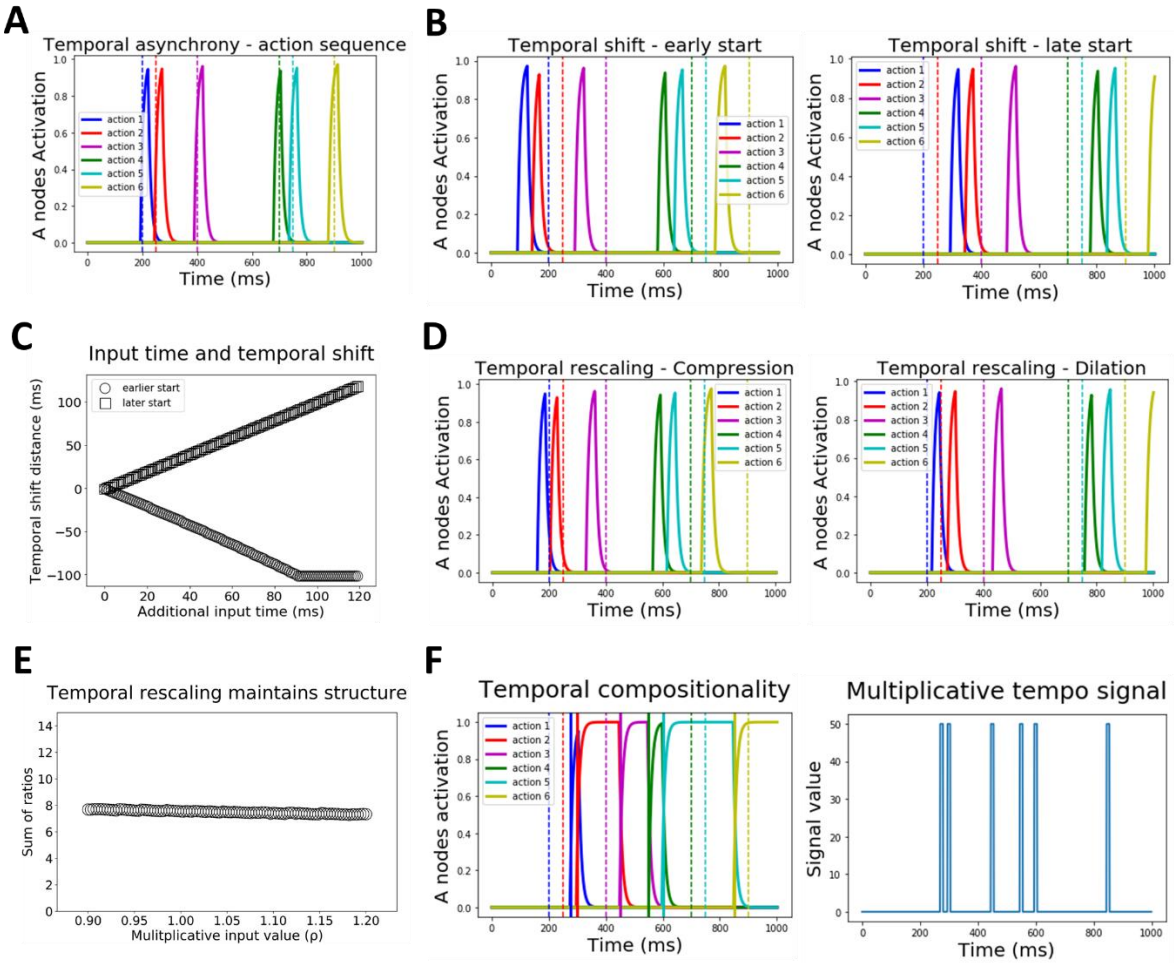

Figure 3. Temporal properties of the ACDC model. A. Simulation 2: Reproduction of action sequence with temporal asynchrony. Each action (i.e. A node activation, color coded) is produced at the precise desired time indicated by the vertical dashed line (also color coded), within a 1 second time window.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

indicating a preserved temporal structure even though the sequence is rescaled. **F. Simulation 5: Temporal compositionality.** The left panel shows how A nodes activity are activated on the tempo described by the multiplicative signal (left panel). Vertical dashed and solid lines on the left panel indicate the timing of each action for the previous and novel tempo respectively. As shown, the respective A nodes become active on the novel tempo.

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 (ρ = 0.1) is sent to the Go node during the period in which sustained motor activation is needed. Figure 4 left shows that applying such a signal between the second and third actions allows motor activation of the second Action node (red trace) to be sustained until the third action is executed (purple trace). Second, via a learning mechanism, the weight value between a specific Action-NoGo node pair can be decreased to induce sustained activation of the Action node (Fig. 4, right). Note that we focus here on the mechanistic properties of the model, rather than proposing how the Action-NoGo weights may be learned to support sustained activation.

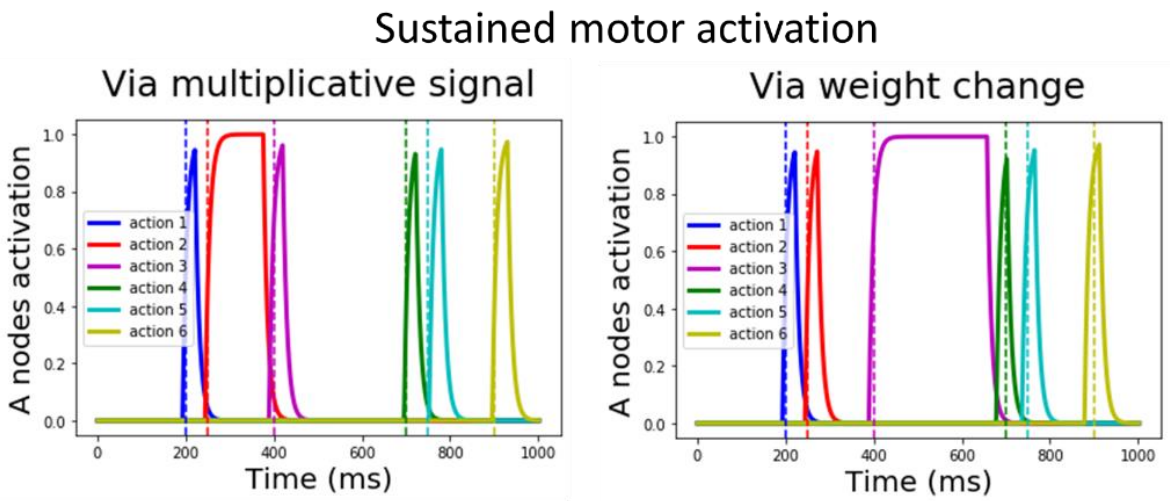

Figure 4. Simulation 6: Sustained motor activation. Both panels demonstrate that the ACDC model is able to output sustained motor activation as desired within a sequence. The left panel shows the results of applying a multiplicative signal (ρ = 0.1) to the second No Go node, inducing a sustained activation 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 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 leveraged temporal compositionality to allow the ACDC model to play the riff but now based on a bossa nova tempo without further training (Fig. 5B-C and Audio file 2). 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).

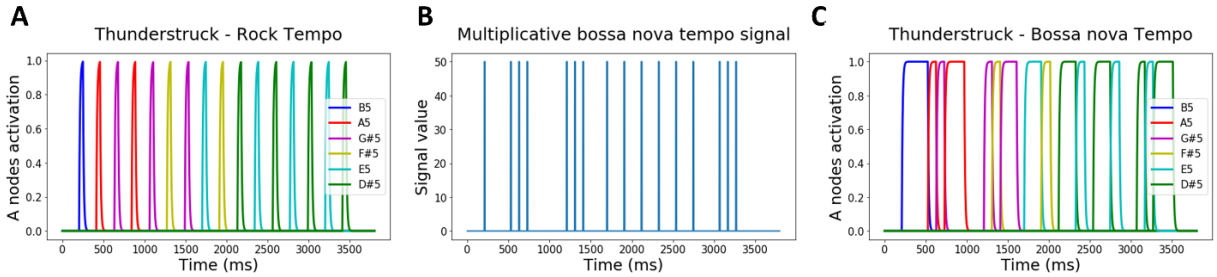

Figure 5. Simulation 7: the ACDC produces the Thunderstruck song. A. Second guitar riff from ACDC's (the group) Thunderstruck song. The riff is composed of 16 sequential actions creating a isosynchronous rock rhythm over a window of 3500 ms (given a 140 bpm tempo). Each action is associated to a color 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 principal component (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[63–66]. In simulation 8, our model learns to produce a single action at distinct interval timings (i.e. 200, 400, 600 and 800 ms). For each timing, the model produces 500 reaction times (RTs), from which we extract the standard deviation (SD), and reproduce this process for 100 simulations and two noise values (gaussian random noise with zero mean and SD of 0.01 or 0.05 is added to the model equations 3-5 and 7). Reproducing empirical patterns, figure 6 shows that the SD of RTs increases as a function of interval timing, and thereby demonstrates that the ACDC model displays scalar variability (see also [67]). Furthermore, the SD value range also increases with noise values. This effect is explained in our model by having a fixed negative bias on the Action nodes in the motor layer. Such a feature reduces to having an accumulation-to-bound process for action execution. Hence, given a specific amount of noise, longer RTs are associated to wider RT distributions (i.e. larger SD, [68]). The underlying reason is that the effect of noise on evidence accumulation is amplified as time elapses.

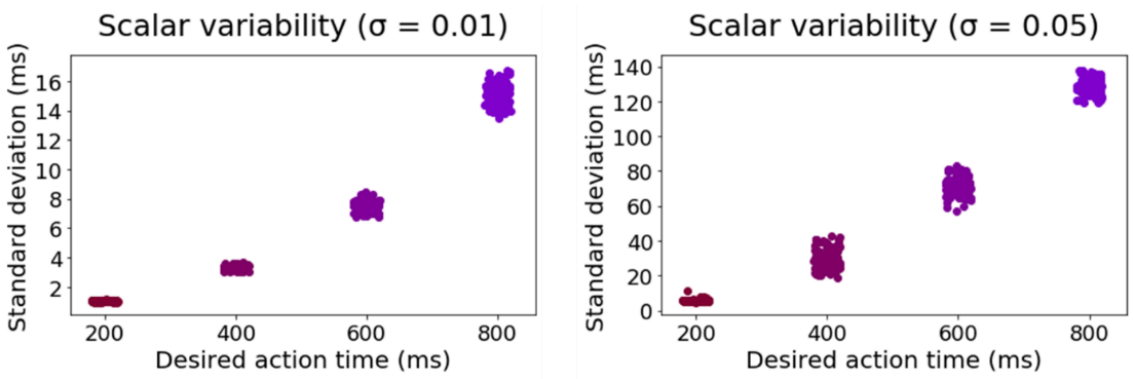

Figure 6. Simulation 8: The ACDC model displays scalar variability. Left (low noise value = 0.01) and right (high noise value = 0.05) panels show that the standard deviation of RTs increases as a function of the desired action time (i.e. interval timing). Moreover, higher noise values increase the range 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[69–71], parietal[72–75], premotor[76] and prefrontal[77–79] cortices whenever an agent has to hold in

working memory task-relevant stimulus features (e.g., spatial location). Theoretical work suggests that persistent activation patterns emerge from recurrently connected networks that settle in one of multiple potential attractor state[80–82]. Second, as motivated in the introduction, sequential activity has also been observed in distinct sequential behaviors such as spatial navigation[8] and bird song[83–86].

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

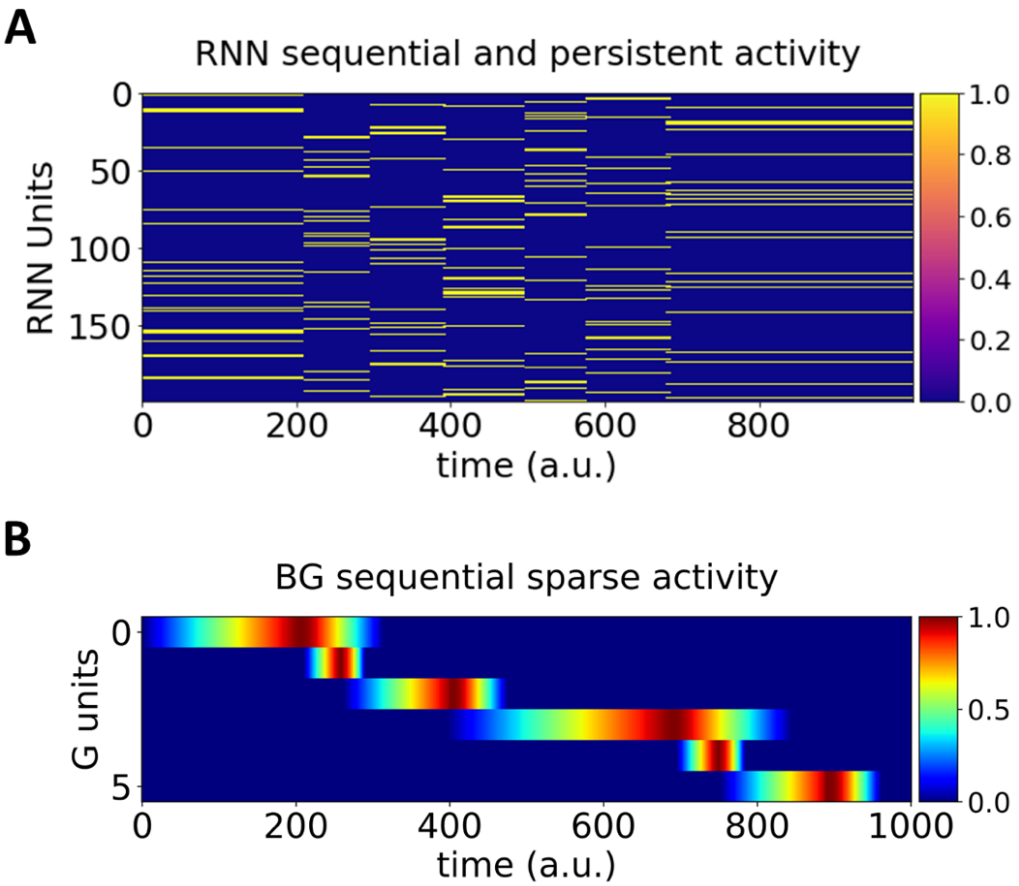

Figure 7. Simulation 9: A. Sequential and persistent activation of clustered neural populations within the RNN. The y-axis represents each RNN unit, the x-axis represents time. The first cluster is activated by the input layer, and maintains persistent activity until the first action is executed. At that moment, via excitatory projections from the Action nodes (Fig. 1C) to the following (i+1th) cluster in the RNN (Fig. 1B) gets activated, and thus displays persistent activation, and so forth via the cortico-basal ganglia loops (light blue arrows in Fig. 1). Color bar represents firing rate. **B**. **Sequential and sparse activation in the BG.** The y-axis represents the G unit activity over time (x-axis). Each G unit responds in a sequential and transient manner, as has been shown in neurophysiological single-cell recordings of the BG (e.g., [4]). Color bar represents normalized firing rate.

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 ith 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[2,8,11] or monkey[1] neurophysiological recordings, which reveal sequential sparse activation (individual neurons display quick and transient activation as behavior unfolds). Notably however, the Go nodes in the BG module of our model display qualitatively similar sequential and sparse activation patterns as that seen empirically in the BG (Fig. 7B; see figures 2A, 3B, 1E, 8C, respectively of [2,3,5,88]).

Model regimes and robustness

A useful model should be robust to variations in its key parameters and/or should exhibit regimes in which the model exhibits qualitatively different features[23,89–91]. To examine the sensitivity of our main results to parameter variations, we start by exploring the model regimes in a network with fixed connectivity. We focus on the recurrent weights within the excitatory RNN units, as well as the weights from the G to A units, which we consider similar to recurrent and feed-forward weights in traditional associative chain models[24]. We explore the necessary weight combinations for the model to execute an entire sequence of 6 actions, within a temporal window of 1 sec (as described in the learning subsection). Figure 8A shows that weight values from G to A units subtend the number of actions that can be executed within the 1 sec temporal window: higher G to A weights progressively lead to sequences with more actions. Notably, recurrent weight values do not influence the model's ability to produce the sequence (no impact of variation along the x-dimension). In contrast to previous models[23], recurrent and G to A weights (i.e., feed-forward weights) do not control the model regimes, i.e., whether the model produces damping, sequential or persistent activity[2].

As shown in simulation 6, one way of controlling persistent activity is through the weight values of the A to N unit projections: activation of a particular action shuts down that action via feedback projections to

[2] This difference emerges from divergent architecture between our and previous models. Persistent activation of each action during the entire sequence cannot emerge in our model. Indeed, action execution is automatically associated to a switch in attractor state within the RNN. This hinders the previous state to continuously activate its G unit and hence cannot provide evidence for the previous A unit. Therefore, if any, persistent activation of an action can be maintained only up until the next action is executed.

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

As noted above, a prominent feature of our model is that action execution triggers switches in RNN attractor states through parallel projections to its excitatory and inhibitory units. For the switch to take place, the gain must be stronger on projections to RNN excitatory vs inhibitory units (i.e. $\gamma_E > \gamma_I$), allowing the previous cluster to shutdown while also activating the next cluster in line. To explore the robustness of this property, we used the learned sequence in simulation 1, and parametrically modulated both gain values (γ_E and γ_I values are represented on the y and x dimension, respectively). As depicted in figure 8C, a vast range of gain combinations allows the model to fully and precisely reproduce the learnt sequence (yellow space), as long as $\gamma_E > \gamma_I$. Note that some combinations allow for a partial reproduction of the sequence.

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

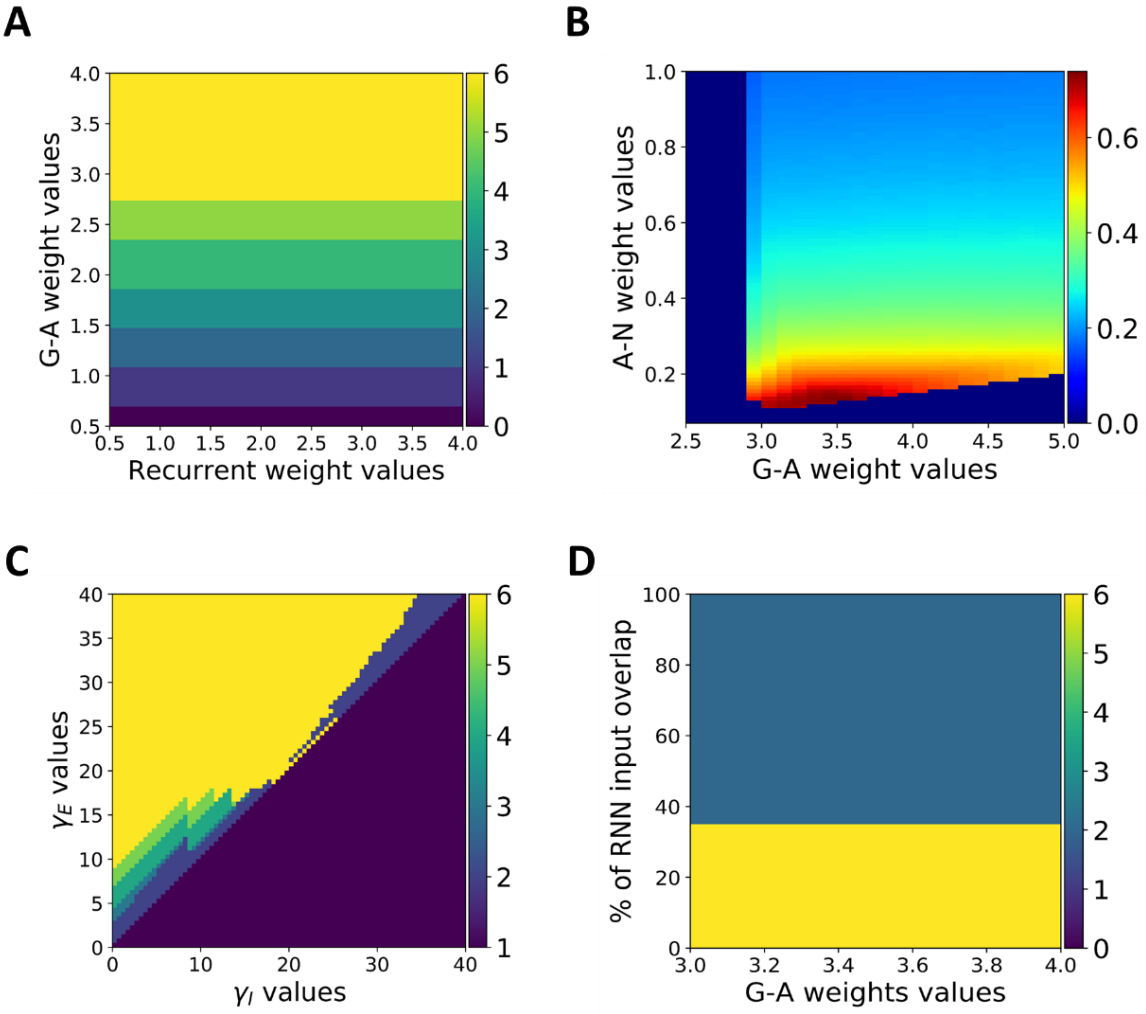

Figure 8. Model regimes and robustness. A. Number of actions within a sequence as a function of feedforward and recurrent weights. The y- and x-axis represent G-A and recurrent weight values, respectively. As depicted, the G-A weight values control whether the model can produce six actions within a 1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

Discussion

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

Encoding order as attractor state switches in the RNN

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

Alternative models have proposed different mechanisms for encoding ordinal position. Some models possess a temporal context layer whose state is modified dynamically as time passes[93–96]. Other models

assume that the network input (used to learn the sequence) is itself sequential in nature[27,28], and learning the spatiotemporal signal depends on the sequential nature of the input. Our model is free of this assumption; the network input is a single pulse of activation, but can nevertheless reproduce a precisely timed spatiotemporal signal. This ability emerges from the feedback loop from thalamic Action nodes to the cortical RNN, triggering transitions to a subsequent cortical attractor. One can therefore consider motor output as part of the teaching input signal to the RNN; because motor activation unfolds sequentially in our model, the sequential nature of the teaching signal emerges from our network architecture.

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

Motor sequence flexibility as inputs to the basal ganglia

Humans can adapt their motor output almost instantaneously given external or internal stimuli. For instance, musicians can modify the tempo of a song upon signaling of the conductor. Such flexibility necessarily needs to stem from fast reconfiguration of neural dynamics, rather than emerge from changes in networks weights[12]. Murray and Escola[28] proposed a model of interconnected medium spiny neurons in the striatum that can apply such dynamic reconfiguration. In particular, their model could perform temporal rescaling of sparse sequential activity. Yet, flexibility in this model is constrained to isosynchronous sequences (see also[67,101]). However, a recent model[22] making use of eligibility traces[102–106], manages to learn precise asynchronous spatiotemporal sequence learning. Still, it is unclear how such a model can rescale asynchronous sequences, and neither of these models is capable of exhibiting temporal compositionality. Nevertheless, ring-like models with synaptic depression[25,90] could potentially account for these properties. Indeed, temporal rescaling in these models is often implemented as changes in the background current (higher current levels lead to faster rescaling). Therefore, to produce asynchronous sequences, one could imagine a dynamical background current which is null when no action has to be provided, resulting in a silent network, and turned on when the sequence has to be resumed. However, given a silent noisy network, reactivation of a global current will induce activation of

576 a random cluster in the sequence and not necessarily the next cluster in line[3], thereby not displaying
577 robustness in sequence production.

578 Crucially, the ACDC model can perform temporal rescaling for both iso-synchronous and asynchronous
579 sequencing, and it can also flexibly switch the tempo altogether through a multiplicative signal to the
580 BG. Our model proposes a more robust approach to sequence flexibility. The sequence in our model is
581 chained by the execution of the previous action, thus our model does not require an explicit
582 representation of the most recent action (see footnote 3). Indeed, what controls the sequence is a series
583 of local signals (i.e., selective feedback from the motor thalamus to the RNN[49]) rather than a global
584 signal, and what controls the timing is a global input to the BG. This feature is key for our model to
585 account for the distinct types of flexibility described in the results.

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

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

[3] Indeed, for such a system to be viable, the network would need an additional piece of information, which is a memory of which cluster (or unit) was last active. The reason is that, in ring-like models the background current is global (i.e., sent to all units). Thus, if the network goes silent, reactivation of the network would not ensure that the appropriate cluster (i.e. the one next in line) becomes active; i.e. noise would randomly activate a cluster and the sequence would restart from that cluster (see implementation of [28]). Consequently, some extra input signal (representing a memory of which cluster was last active) should give an advantage to the next cluster in line, in order to ensure that sequence order is maintained. This extra input signal could take the form of recurrent connections within the cluster, controlling the activation decay of that cluster such that it would still be active after long periods of time (as opposed to all other clusters that would be fully silent; see [22]). Note however that recurrent connections would put a break on the synaptic depression, as these two parameters trade off; therefore, formal implementation of this proposal would be needed to understand the model regimes.

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

612 ***Limitations and future directions***

613 As previously noted, some of the implementational and biological details of our model remain to be
614 worked out. First, we simplified the BG gating circuitry, to focus on the G and N populations,
615 summarizing their effects on downstream thalamus but omitting the disinhibitory circuitry involving the
616 substantia nigra. Many previous models, including our own, have simulated the more complete direct
617 and indirect pathways but we did not feel this detail was necessary for the present purposes. Second,
618 reinforcement learning of action timing is conceptually thought to take the form of a three-factor
619 hebbian learning rule[43,103,137–139], where neurons subtending a rewarding behavior (and hence forming a
620 specific cortical activity patterns) increase their connectivity to D1-receptor containing striatal "Go"
621 populations via dopaminergic activity bursts stemming from midbrain nuclei[90,140–143]. While we do not
622 challenge this mechanism, we focused learning in our model to one synapse downstream, from Go
623 nodes to thalamic motor neurons. Indeed, recent evidence suggest that error-driven learning can be
624 achieved via manipulation of BG outputs to thalamus[144]. Third, although much evidence indicates that
625 the BG learn via reinforcement learning[145–147] (i.e., depending on whether rewards are better or worse
626 than expected), we incorporated a signed error in our learning mechanism which is more powerful for
627 timing signals. The supervised delta rule is only used at the output of our network to optimize readout
628 weights. Thus, we do not view this learning rule as implausible in our model given that we focus on

629 sequence learning situations in which signed error feedback is provided (see [61]), such as when a tutor
630 teaches you how to play the drums and holds the tempo, or in bird-song learning[84] (where a tutor is
631 available to provide signed error). There are several biologically plausible implementations of the delta
632 rule when such error signals are available (e.g., [148,149]). Our learning rule in the BG-thalamus thus
633 summarizes the contributions of these systems in conjunction. Nonetheless, future work should
634 investigate how and whether complex precise sequences may emerge solely based on reinforcement
635 learning[4]. In contrast to ACDC, other models have directly trained the entire time-series of individual
636 RNN units to match empirical data[38,150], using continuous learning signals (also see[34,36,37,39]). Hence, the
637 neurophysiological simulations in ACDC (i.e., sequential sparse activation and attractor states) emerge
638 only from the proposed theoretical architecture in the context of behavioral experience. Fourth, we
639 implemented a single inhibitory neuron in the RNN-PMC module. Our focus was on the role of the
640 functional role of inhibitory neurons on the transition of attractor states[28], and certainly this single unit
641 could be replaced by a larger population. Future versions of the model should include a broader pool of
642 inhibitory neurons in the RNN as they have been shown to exhibit mixed selectivity to multiple aspects
643 of a task[151]. Fifth, one assumption in our model is orthogonality of all projections to the RNN. Although
644 we showed that this assumption can be relaxed to a certain extent, this feature ensures there is no
645 ordinal interference during action sequence execution. In the brain, this orthogonality may be
646 implemented via mixed selectivity of excitatory frontal neurons that ensure downstream readouts
647 without interence[152,153]. Interestingly, Márton et al.[154] recently developed a RNN model of cortico-
648 striatal interactions optimized to learn oculomotor sequences. Similar sequences were performed by
649 awake monkeys while activity was recorded in their dorsolateral prefrontal and striatal areas. Learning
650 to implement the correct actions for each sequence pulled apart the representational structure of action
651 sequences in activity space both in the model and neuronal recordings. Whereas ordinal representations
652 in our network were hardwired as orthogonal vectors in the RNN (in order to avoid interference), the
653 work of Márton et al. suggests this may emerge naturally through learning.

654

655 Our model simulates action sequences such as those needed to play the guitar or the piano. Within this
656 context, each action is represented as a discrete entity. However, many daily life action sequences are
657 subtended by more continuous actions, as for instance when playing violin with a bow. The ACDC could

[4] In principle, the ACDC model could learn sequences solely based on RL. However, learning would be much more tedious[43].

658 be expanded by having more continuous representations of action plans and execution in our BG-
659 thalamus module. Based on dynamic field theory, one potential approach would be to represent actions
660 as dynamic neural fields[155–157], which have been shown to successfully model more continuous reaching
661 actions[158]. Moreover, these continuous action representations in the BG may require additional inputs
662 from the cerebellum for movement coordination [159] or sequence prediction for motor control[160].
663 Moreover, our model (as others [22,28,67]) was specifically engineered to account for spatiotemporal
664 sequences and how these may be flexibly manipulated. This in contrast to other instantiations[35,161] of
665 RNNs (i.e. reservoir computing) that find natural solutions to diverse tasks involving distinct
666 psychological processes (e.g., memory, time estimation, decision-making).

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

672

673 Methods

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

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

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

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

681 Crucially, the dynamics within the ACDC model evolve as a sequential unfolding of RNN-BG-thalamus-
682 RNN (i.e., cortico-basal ganglia) loops, depicted by the light blue arrows in figure 1. The sequence starts
683 with the activation of a cluster (i.e., densely interconnected) of excitatory RNN neurons (Fig. 1B). Each
684 cluster will come to encode the ith element in the action sequence. As opposed to single unit, clustered
685 neurons provide a biologically plausible mechanism for supporting persistent activation within the

686 cluster given a phasic input (i.e., an attractor, [81,166]). In prefrontal cortical – BG models, such clusters are
687 referred to as "stripes" based on their anatomical existence, and are independently gated by BG[167].
688 Once a cluster is activated, the RNN temporarily settles on an attractor state indicating the ordinal
689 position (order or rank) in the sequence, analogous to how distinct PFC stripes code for ordinal positions
690 in phonological loop tasks[167]. However, in ACDC such clusters emerge naturally via learning rather than
691 hard-coded anatomical entities. Moreover, attractor states are maintained via a specific ratio of
692 excitatory to inhibitory inputs: each excitatory neuron projects to a common single inhibitory neuron
693 (orange circle in Fig. 1B) which reciprocally inhibits all excitatory RNN neurons. As long as the ratio of
694 excitatory to inhibitory inputs is not perturbed by another input (see below), activation in the cluster will
695 persist and the RNN will continue representing the ith order in the sequence.

696 In turn, each excitatory RNN cluster projects to its corresponding "Go" unit in the BG (blue arrow 1 from
697 i^{th} cluster in Fig. 1B to G node in Fig. 1C), and each Go cell accumulates evidence for the jth action
698 associated to the ith order (see [134,168] for related computational models of evidence accumulation in
699 these units, and [169] for empirical data). Striatal Go cells, via the basal ganglia direct pathway
700 machinery[170,171], facilitate response execution by projecting towards the corresponding motor thalamus
701 neurons, from here on termed Action nodes for simplicity (blue arrow 2 from Go to Action nodes in Fig.
702 1C). The BG component summarizes the contributions of more detailed BG circuitry[140,167,172]. In these
703 models striatal neurons accumulate evidence, which via the direct and indirect pathways leads to
704 categorically discrete signals in BG output nuclei, and to disinhibition of the thalamus (e.g., [132,168]). These
705 patterns are also observed empirically in terms of striatum accumulation signals and discrete
706 downstream responses in BG output nuclei once a threshold of accumulation is reached[169]. Here, we
707 lumped together the double inhibition from striatum to Globus Pallidus (GP) and from GP to the
708 thalamus into a single excitatory projection to keep the model simple and tractable. Interestingly,
709 optogenetic stimulation of the GP has been shown to increase the firing rate of motor thalamus
710 neurons[173].

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

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

731 Several features of the model should be highlighted. First, each cluster activation within the RNN acts as
732 an attractor state representing the ith element in the sequence. Interestingly, cells in the monkey PMC
733 code for the position in sequence, regardless of the actual movement produced during that position[178–
734 184]. We therefore assume that the neurons forming each cluster represent rank-order-selective neurons
735 whose activation unfolds sequentially: the RNN encodes order information.

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

744 Third, as in many cortico-BG models (e.g., [134,190]), and motivated by anatomical data[191] our model is
745 characterized by topographical organization of actions across the BG circuit and its outputs (i.e., indexed
746 in our model by the subscript j associated in the G-A-N triplet projections). Recent evidence further
747 confirms topographical action representations in BG-thalamocortical loops[192–194], whereby causal

activation of specific subregions is related to specific output behaviors[195], and is also supported by human neuroimaging[196] and monkey/rodent neurophysiology studies[1,197–202]. However, in contrast to previous models in which BG gating affords action selection of the corresponding cortical action, in the ACDC model BG gating triggers a cortical dynamical state that initiates the evolution of the *subsequent* item in the sequence.

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

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

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

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

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

equation 1

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

Second, 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, and is described in equation 2:

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

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

Mathematical description of the model dynamics

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

The dynamics within the ACDC model represent the sequential unfolding of RNN-BG-thalamus-RNN (i.e., cortico-basal ganglia) loops, depicted by the light blue arrows in 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. γ_E is the gain on Action nodes activation projected to the excitatory RNN neurons (see below for the functional property of this parameter). Θ, the non-linear transformation function, is governed by $\Theta(x) = (2 / (1 + e^{-\lambda x})) - 1$ (where λ is the gain parameter and with additional non-linearity at zero, i.e. $\Theta(x) = 0$ if $\Theta(x) < 0$); and τ_{rnn} is the encoding constant. Note that input projections and all Action nodes to RNN projections are orthogonal (i.e. some RNN excitatory neurons receive inputs from the input layer, whereas others receive from inputs from Action nodes; each projection excites 20 RNN units). The shared inhibitory x_I activation is described by 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 γ_I respectively represent the weights from the excitatory RNN neurons to their shared inhibitory neuron, the weights from the Action nodes to the shared inhibitory neuron, and the gain on Action nodes activation for the projections towards the inhibitory neuron in the RNN.

In turn, each excitatory RNN cluster projects to its corresponding "Go" cell in the BG (blue arrow 1 from Fig. 1B to Go node in Fig. 1C), and each Go cell accumulates evidence for the jth action associated to the ith order, following 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 jth Go units, W_{ij} is the weight matrix representing connectivity between RNN and Go units, x_i is the acitivity of the RNN excitatory units, J^{GN} is the inhibitory weight between the jth No Go and Go nodes, n_j is the activation of the jth No Go node, and τ_g is the encoding constant (with $\tau_g >>> 0$, thereby simulating evidence accumulation-like dynamics). Non-linearity at zero is also applied to Go-units.

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

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

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
b	0.5
τ_w	2
η	0.4
ϕ	0.01
J^{IE}	0.1

τ_a, τ_n	10
γ_E / γ_I	21.4 / 21
τ_{rnn}, x^{in}, J^{EI}, J^{EA}, J^{IA}, J^{GN}, J^{NA}	1
λ_{rnn} / λ_a	10 / 10000
τ_g	1000

References

1. Jin, D. Z., Fujii, N. & Graybiel, A. M. Neural representation of time in cortico-basal ganglia circuits. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19156–19161 (2009).

2. Mello, G. B. M., Soares, S. & Paton, J. J. A scalable population code for time in the striatum. *Curr. Biol.* **25**, 1113–1122 (2015).

3. Dhawale, A. K. *et al.* Automated long-Term recording and analysis of neural activity in behaving animals. *Elife* **6**, 1–40 (2017).

4. Gouvêa, T. S. *et al.* Striatal dynamics explain duration judgments. *Elife* **4**, 1–14 (2015).

5. Bakhurin, K. I. *et al.* Differential Encoding of Time by Prefrontal and Striatal Network Dynamics. *J. Neurosci.* **37**, 854–870 (2017).

6. Pastalkova, E., Itskov, V., Amarasingham, A. & Buzsaki, G. Internally Generated Cell Assembly Sequences in the Rat Hippocampus. *Science (80-.).* **321**, 1322–1327 (2008).

7. MacDonald, C. J. & Tonegawa, S. Crucial role for CA2 inputs in the sequential organization of CA1 time cells supporting memory. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).

8. Eichenbaum, H. Time cells in the hippocampus: A new dimension for mapping memories. *Nat. Rev. Neurosci.* **15**, 732–744 (2014).

9. Nicola, W. & Clopath, C. A diversity of interneurons and Hebbian plasticity facilitate rapid compressible learning in the hippocampus. *Nat. Neurosci.* **22**, 1168–1181 (2019).

10. Luczak, A., Barthó, P., Marguet, S. L., Buzsáki, G. & Harris, K. D. Sequential structure of neocortical spontaneous activity in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 347–352 (2007).

11. Harvey, C. D., Coen, P. & Tank, D. W. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484**, 62–68 (2012).

12. Remington, E. D., Egger, S. W., Narain, D., Wang, J. & Jazayeri, M. A Dynamical Systems Perspective on Flexible Motor

Timing. *Trends Cogn. Sci.* **22**, 938–952 (2018).

13. Sompolinsky, H. & Kanter, I. Temporal association in asymmetric neural networks. *Phys. Rev. Lett.* **57**, 2861–2864 (1986).

14. Kleinfeld, D. Sequential state generation by model neural networks. *Proc. Natl. Acad. Sci.* **83**, 9469–9473 (1986).

15. Herrmann, M., Hertz, J. A. & Prügel-Bennett, A. Network : Computation in Neural Systems Analysis of synfire chains Analysis of synfire chains. *Netw. Comput. Ned Syst.* **6**, 403–414 (1995).

16. Graybiel, A. M. The Basal Ganglia and Chunking of Action Repertoires. *Neurobiol. Learn. Mem.* **70**, 119–136 (1998).

17. Geddes, C. E., Li, H. & Jin, X. Optogenetic Editing Reveals the Hierarchical Organization of Learned Action Sequences. *Cell* **174**, 32-43.e15 (2018).

18. Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L. & Haith, A. M. Motor learning. *Compr. Physiol.* **9**, 613–663 (2019).

19. Fiete, I. R., Senn, W., Wang, C. Z. H. & Hahnloser, R. H. R. Spike-Time-Dependent Plasticity and Heterosynaptic Competition Organize Networks to Produce Long Scale-Free Sequences of Neural Activity. *Neuron* **65**, 563–576 (2010).

20. Dlesmann, M., Gewaltig, M. O. & Aertsen, A. Stable propagation of synchronous spiking in cortical neural networks. *Nature* **402**, 529–533 (1999).

21. Abeles, M. *Corticonics: Neural circuits of the cerebral cortex*. (Cambridge University Press, 1991).

22. Cone, I. & Shouval, H. Z. Learning precise spatiotemporal sequences via biophysically realistic learning rules in a modular, spiking network. *Elife* **10**, 2020.04.17.046862 (2021).

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

24. Veliz-Cuba, A., Shouval, H. Z., Josić, K. & Kilpatrick, Z. P. Networks that learn the precise timing of event sequences. *J. Comput. Neurosci.* **39**, 235–254 (2015).

25. Goodbody, S. J. & Wolpert, D. M. Temporal and amplitude generalization in motor learning. *J. Neurophysiol.* **79**, 1825–1838 (1998).

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

27. Maes, A., Barahona, M. & Clopath, C. Learning spatiotemporal signals using a recurrent spiking network that discretizes time. *PLoS Comput. Biol.* **16**, 1–26 (2020).

28. Murray, J. M. & Escola, G. S. Learning multiple variable-speed sequences in striatum via cortical tutoring. *Elife* **6**, 1–24 (2017).

29. Litwin-Kumar, A. & Doiron, B. Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nat. Commun.* **5**, 1–12 (2014).

30. Zenke, F., Agnes, E. J. & Gerstner, W. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nat. Commun.* **6**, 1–13 (2015).

31. Del Giudice, P., Fusi, S. & Mattia, M. Modelling the formation of working memory with networks of integrate-and-fire neurons connected by plastic synapses. *J. Physiol. Paris* **97**, 659–681 (2003).

32. Mongillo, G., Amit, D. J. & Brunel, N. Retrospective and prospective persistent activity induced by Hebbian learning in a recurrent cortical network. *Eur. J. Neurosci.* **18**, 2011–2024 (2003).

33. Gillett, M., Pereira, U. & Brunel, N. Characteristics of sequential activity in networks with temporally asymmetric Hebbian learning. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 29948–29958 (2020).

34. Hardy, N. F., Goudar, V., Romero-Sosa, J. L. & Buonomano, D. V. A model of temporal scaling correctly predicts that motor timing improves with speed. *Nat. Commun.* **9**, 1–14 (2018).

35. Orhan, A. E. & Ma, W. J. A diverse range of factors affect the nature of neural representations underlying short-term memory. *Nat. Neurosci.* **22**, 275–283 (2019).

36. Laje, R. & Buonomano, D. V. Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nat. Neurosci.* **16**, 925–933 (2013).

37. Buonomano, D. V. Harnessing Chaos in Recurrent Neural Networks. *Neuron* **63**, 423–425 (2009).

38. Rajan, K., Harvey, C. D. D. & Tank, D. W. W. Recurrent Network Models of Sequence Generation and Memory. *Neuron* **90**, 128–142 (2016).

39. Goudar, V. & Buonomano, D. V. Encoding sensory and motor patterns as time-invariant trajectories in recurrent neural networks. *Elife* **7**, 1–28 (2018).

40. Logiaco, L., Abbott, L. F. & Escola, S. Thalamic control of cortical dynamics in a model of flexible motor sequencing. *Cell Rep.* **35**, 109090 (2021).

41. Sussillo, D. & Abbott, L. F. Generating Coherent Patterns of Activity from Chaotic Neural Networks. *Neuron* **63**, 544–557 (2009).

42. Bellec, G. *et al.* A solution to the learning dilemma for recurrent networks of spiking neurons. *Nat. Commun.* **11**, 1–15 (2020).

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

44. Murray, J. M. Local online learning in recurrent networks with random feedback. *Elife* **8**, 1–25 (2019).

45. Wang, J., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* **21**, 102–112 (2018).

46. Recanatesi, S., Pereira-Obilinovic, U., Murakami, M., Mainen, Z. & Mazzucato, L. Metastable attractors explain the

932 variable timing of stable behavioral action sequences. *Neuron* 1–15 (2021). doi:10.1016/j.neuron.2021.10.011

933 47. Franklin, N. T. & Frank, M. J. Compositional clustering in task structure learning. *PLOS Comput. Biol.* **14**, e1006116
934 (2018).

935 48. Franklin, N. T. & Frank, M. J. *Generalizing to generalize: Humans flexibly switch between compositional and conjunctive*
936 *structures during reinforcement learning*. *PLoS Computational Biology* **16**, (2020).

937 49. Burke, D. A., Rotstein, H. G. & Alvarez, V. A. Striatal Local Circuitry: A New Framework for Lateral Inhibition. *Neuron* **96**,
938 267–284 (2017).

939 50. Halassa, M. M. & Sherman, S. M. Thalamocortical Circuit Motifs: A General Framework. *Neuron* **103**, 762–770 (2019).

940 51. Rikhye, R. V., Wimmer, R. D. & Halassa, M. M. Toward an Integrative Theory of Thalamic Function. *Annu. Rev. Neurosci.*
941 **41**, 163–183 (2018).

942 52. Rikhye, R. V., Gilra, A. & Halassa, M. M. Thalamic regulation of switching between cortical representations enables
943 cognitive flexibility. *Nat. Neurosci.* **21**, 1753–1763 (2018).

944 53. Wymbs, N. F., Bassett, D. S., Mucha, P. J., Porter, M. A. & Grafton, S. T. Article Differential Recruitment of the
945 Sensorimotor Putamen and Frontoparietal Cortex during Motor Chunking in Humans. *Neuron* **74**, 936–946 (2012).

946 54. Lungu, O. *et al.* Striatal and hippocampal involvement in motor sequence chunking depends on the learning strategy.
947 *PLoS One* **9**, 25–27 (2014).

948 55. Doyon, J., Gabitov, E., Vahdat, S., Lungu, O. & Boutin, A. ScienceDirect Current issues related to motor sequence
949 learning in humans. *Curr. Opin. Behav. Sci.* **20**, 89–97 (2017).

950 56. Graybiel, A. M. & Grafton, S. T. The striatum: Where skills and habits meet. *Cold Spring Harb. Perspect. Biol.* **7**, 1–14
951 (2015).

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

954 58. Abrahamse, E. L., Ruitenberg, M. F. L., de Kleine, E. & Verwey, W. B. Control of automated behavior: Insights from the
955 discrete sequence production task. *Front. Hum. Neurosci.* **7**, 1–16 (2013).

956 59. Diedrichsen, J. & Kornysheva, K. Motor skill learning between selection and execution. *Trends Cogn. Sci.* **19**, 227–233
957 (2015).

958 60. Verwey, W. B., Shea, C. H. & Wright, D. L. A cognitive framework for explaining serial processing and sequence
959 execution strategies. *Psychon. Bull. Rev.* **22**, 54–77 (2014).

960 61. Kornysheva, K. *et al.* Neural Competitive Queuing of Ordinal Structure Underlies Skilled Sequential Action. *Neuron* **101**,
961 1166-1180.e3 (2019).

962 62. Forstmann, B. U. *et al.* Striatum and pre-SMA facilitate decision-making under time pressure. *Proc. Natl. Acad. Sci. U. S.*

963 *A.* **105**, 17538–17542 (2008).

964 63. Acerbi, L., Wolpert, D. M. & Vijayakumar, S. Internal Representations of Temporal Statistics and Feedback Calibrate
965 Motor-Sensory Interval Timing. *PLoS Comput. Biol.* **8**, (2012).

966 64. Rakitin, B. C. *et al.* Scalar expectancy theory and peak-interval timing in humans. *J. Exp. Psychol. Anim. Behav. Process.*
967 **24**, 15–33 (1998).

968 65. Jazayeri, M. & Shadlen, M. N. Temporal context calibrates interval timing. *Nat. Neurosci.* **13**, 1020–1026 (2010).

969 66. Ivry, R. B. & Hazeltine, R. E. Perception and production of temporal intervals across a range of durations: Evidence for a
970 common timing mechanism. *J. Exp. Psychol. Hum. Percept. Perform.* **21**, 3–18 (1995).

971 67. Egger, S. W., Le, N. M. & Jazayeri, M. A neural circuit model for human sensorimotor timing. *Nat. Commun.* **11**, 1–14
972 (2020).

973 68. Ratcliff, R. & Rouder, J. N. Modelling response times for two choice decisions. *Psychol. Sci.* **9**, 347–356 (1998).

974 69. Miyashita, Y. & Chang, H. S. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature*
975 **331**, 68–70 (1988).

976 70. Nakamura, K. & Kubota, K. Mnemonic firing of neurons in the monkey temporal pole during a visual recognition
977 memory task. *J. Neurophysiol.* **74**, 162–178 (1995).

978 71. Erickson, C. A. & Desimone, R. Responses of macaque perirhinal neurons during and after visual stimulus association
979 learning. *J. Neurosci.* **19**, 10404–10416 (1999).

980 72. Koch, K. W. & Fuster, J. M. Unit activity in monkey parietal cortex related to haptic perception and temporary memory.
981 *Exp. Brain Res.* (1989). doi:10.1007/BF00247889

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

984 74. Gail, A. & Andersen, R. a. Neural dynamics in monkey parietal reach region reflect context-specific sensorimotor
985 transformations. *J. Neurosci.* **26**, 9376–9384 (2006).

986 75. Klaes, C., Westendorff, S., Chakrabarti, S. & Gail, A. Choosing goals, not rules: deciding among rule-based action plans.
987 *Neuron* **70**, 536–48 (2011).

988 76. Cisek, P. & Kalaska, J. F. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple
989 direction choices and final selection of action. *Neuron* **45**, 801–14 (2005).

990 77. Funahashi, S., Bruce, C. J. & Goldman-Rakic, P. S. Visuospatial coding in primate prefrontal neurons revealed by
991 oculomotor paradigms. *J. Neurophysiol.* **63**, 814–831 (1990).

992 78. Funahashi, S., Bruce, C. J. & Goldman-Rakic, P. S. Mnemonic coding of visual space in the monkey's dorsolateral
993 prefrontal cortex. *J. Neurophysiol.* **61**, 331–349 (1989).

994 79. Miller, E. K., Erickson, C. A. & Desimone, R. Neural mechanisms of visual working memory in prefrontal cortex of the
995 macaque. *J. Neurosci.* **16**, 5154–5167 (1996).

996 80. Wang, X. J. Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* **24**, 455–463 (2001).

997 81. Durstewitz, D., Seamans, J. K. & Sejnowski, T. J. Neurocomputational Models of Working Memory. *Nat. Neurosci.* **3**,
998 1184–1191 (2000).

999 82. Brunel, N. Dynamics and Plasticity of Stimulus-selective Persistent Activity in Cortical Network Models. *Cereb. Cortex*
1000 **13**, 1151–1161 (2003).

1001 83. Hahnloser, R. H. R., Kozhevnikov, A. A. & Fee, M. S. An ultra-sparse code underlies the generation of neural sequences
1002 in a songbird. *Nature* **419**, 65–70 (2002).

1003 84. Okubo, T. S., Mackevicius, E. L., Payne, H. L., Lynch, G. F. & Fee, M. S. Growth and splitting of neural sequences in
1004 songbird vocal development. *Nature* **528**, 352–357 (2015).

1005 85. Kozhevnikov, A. A. & Fee, M. S. Singing-related activity of identified HVC neurons in the zebra finch. *J. Neurophysiol.* **97**,
1006 4271–4283 (2007).

1007 86. Amador, A., Perl, Y. S., Mindlin, G. B. & Margoliash, D. Elemental gesture dynamics are encoded by song premotor
1008 cortical neurons. *Nature* **495**, 59–64 (2013).

1009 87. Recanatesi, S., Pereira, U., Murakami, M., Mainen, Z. & Mazzucato, L. Metastable attractors explain the variable timing
1010 of stable behavioral action sequences. *bioRxiv* 0–40 (2020). doi:10.1101/2020.01.24.919217

1011 88. Rueda-Orozco, P. E. & Robbe, D. The striatum multiplexes contextual and kinematic information to constrain motor
1012 habits execution. *Nat. Neurosci.* **18**, 453–462 (2015).

1013 89. Verguts, T., Vassena, E. & Silvetti, M. Adaptive effort investment in cognitive and physical tasks: a neurocomputational
1014 model. *Front. Behav. Neurosci.* **9**, 57 (2015).

1015 90. Collins, A. G. E. & Frank, M. J. Cognitive control over learning: creating, clustering, and generalizing task-set structure.
1016 *Psychol. Rev.* **120**, 190–229 (2013).

1017 91. Calderon, C. B., Gevers, W. & Verguts, T. The unfolding action model of initiation times, movement times, and
1018 movement paths. *Psychol. Rev.* 1–65 (2018). doi:10.1037/rev0000110

1019 92. Ju, H. & Bassett, D. S. Dynamic representations in networked neural systems. *Nat. Neurosci.* **23**, 908–917 (2020).

1020 93. Burgess, N. & Hitch, G. J. A revised model of short-term memory and long-term learning of verbal sequences. *J. Mem.*
1021 *Lang.* **55**, 627–652 (2006).

1022 94. Burgess, N. & Hitch, G. J. Memory for serial order: A network model of the phonological loop and its timing. *Psychol.*
1023 *Rev.* **106**, 551–581 (1999).

1024 95. Page, M. P. A. & Norris, D. The primacy model: A new model of immediate serial recall. *Psychol. Rev.* **105**, 761–781

(1998).

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

97. Kawai, R. *et al.* Motor Cortex Is Required for Learning but Not for Executing a Motor Skill. *Neuron* **86**, 800–812 (2015).

98. Otchy, T. M. *et al.* Acute off-target effects of neural circuit manipulations. *Nature* **528**, 358–363 (2015).

99. Koziol, L. F. & Budding, D. E. *Subcortical structures and cognition: Implications for neuropsychological assessment*. (Springer Science & Business Media., 2009).

100. Collins, A. G. E. & Frank, M. J. Motor Demands Constrain Cognitive Rule Structures. *PLoS Comput. Biol.* **12**, 1–17 (2016).

101. Kozachkov, L. & Michmizos, K. P. Sequence Learning in Associative Neuronal-Astrocytic Networks. in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **12241 LNAI**, 349–360 (2020).

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

103. Izhikevich, E. M. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex* **17**, 2443–2452 (2007).

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

105. Soltoggio, A. & Steil, J. J. Solving the distal reward problem with rare correlations. *Neural Comput.* **25**, 940–978 (2013).

106. Bellec, G. *et al.* Biologically inspired alternatives to backpropagation through time for learning in recurrent neural nets. *arXiv* 1–37 (2019).

107. Sussillo, D., Churchland, M. M., Kaufman, M. T. & Shenoy, K. V. A neural network that finds a naturalistic solution for the production of muscle activity. *Nat. Neurosci.* **18**, 1025–1033 (2015).

108. Russo, A. A. *et al.* Motor Cortex Embeds Muscle-like Commands in an Untangled Population Response. *Neuron* **97**, 953-966.e8 (2018).

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

110. Beierholm, U. *et al.* Dopamine modulates reward-related vigor. *Neuropsychopharmacology* **38**, 1495–1503 (2013).

111. Panigrahi, B. *et al.* Dopamine Is Required for the Neural Representation and Control of Movement Vigor. *Cell* **162**, 1418–1430 (2015).

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

1056 113. Berke, J. D. What does dopamine mean? *Nat. Neurosci.* **21**, 787–793 (2018).

1057 114. Hamid, A. A. *et al.* Mesolimbic dopamine signals the value of work. *Nat. Neurosci.* **19**, 117–126 (2015).

1058 115. Hamid, A. A., Frank, M. J. & Moore, C. I. Dopamine waves as a mechanism for spatiotemporal credit assignment. *Cell*
1059 (2021).

1060 116. Gaidica, M., Hurst, A., Cyr, C. & Leventhal, D. K. Distinct populations of motor thalamic neurons encode action
1061 initiation, action selection, and movement vigor. *J. Neurosci.* **38**, 6563–6573 (2018).

1062 117. Sedaghat-Nejad, E., Herzfeld, D. J. & Shadmehr, R. Reward prediction error modulates saccade vigor. *J. Neurosci.* **39**,
1063 5010–5017 (2019).

1064 118. Augustin, S. M., Loewinger, G. C., O'Neal, T. J., Kravitz, A. V. & Lovinger, D. M. Dopamine D2 receptor signaling on
1065 iMSNs is required for initiation and vigor of learned actions. *Neuropsychopharmacology* **45**, 2087–2097 (2020).

1066 119. van Gaalen, M. M., van Koten, R., Schoffelmeer, A. N. M. & Vanderschuren, L. J. M. J. Critical Involvement of
1067 Dopaminergic Neurotransmission in Impulsive Decision Making. *Biol. Psychiatry* **60**, 66–73 (2006).

1068 120. Pattij, T. & Vanderschuren, L. J. M. J. The neuropharmacology of impulsive behaviour. *Trends Pharmacol. Sci.* **29**, 192–
1069 199 (2008).

1070 121. Buckholtz, J. W. *et al.* Dopaminergic network differences in human impulsivity. *Science (80-.).* **329**, 532 (2010).

1071 122. Pine, A., Shiner, T., Seymour, B. & Dolan, R. J. Dopamine, time, and impulsivity in humans. *J. Neurosci.* **30**, 8888–8896
1072 (2010).

1073 123. Dalley, J. W. & Roiser, J. P. Dopamine, serotonin and impulsivity. *Neuroscience* **215**, 42–58 (2012).

1074 124. Economidou, D., Theobald, D. E. H., Robbins, T. W., Everitt, B. J. & Dalley, J. W. Norepinephrine and dopamine
1075 modulate impulsivity on the five-choice serial reaction time task through opponent actions in the shell and core sub-
1076 regions of the nucleus accumbens. *Neuropsychopharmacology* **37**, 2057–2066 (2012).

1077 125. Frank, M. J., Samanta, J., Moustafa, A. A. & Sherman, S. J. Hold your horses: impulsivity, deep brain stimulation, and
1078 medication in parkinsonism. *Science* **318**, 1309–12 (2007).

1079 126. Lake, J. I. & Meck, W. H. Differential effects of amphetamine and haloperidol on temporal reproduction: Dopaminergic
1080 regulation of attention and clock speed. *Neuropsychologia* **51**, 284–292 (2013).

1081 127. Ratcliff, R. A theory of memory retrieval. *Psychol. Rev.* **85**, 59–108 (1978).

1082 128. Beste, C. *et al.* Dopamine modulates the efficiency of sensory evidence accumulation during perceptual decision
1083 making. *Int. J. Neuropsychopharmacol.* **21**, 649–655 (2018).

1084 129. Yousif, N. *et al.* Dopamine activation preserves visual motion perception despite noise interference of human v5/mt. *J.*
1085 *Neurosci.* **36**, 9303–9312 (2016).

1086 130. Westbrook, A. *et al.* Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work. *Science*

1087 *(80-.).* **367**, 1362–1366 (2020).

1088 131. Heitz, R. P. The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Front. Neurosci.* **8**, 1–19
1089 (2014).

1090 132. Wiecki, T. V & Frank, M. J. A computational model of inhibitory control in frontal cortex and basal ganglia. *Psychol. Rev.*
1091 **120**, 329–355 (2013).

1092 133. Lloyd, K. & Dayan, P. Tamping Ramping: Algorithmic, Implementational, and Computational Explanations of Phasic
1093 Dopamine Signals in the Accumbens. *PLoS Comput. Biol.* **11**, 1–34 (2015).

1094 134. Frank, M. J. Hold your horses: a dynamic computational role for the subthalamic nucleus in decision making. *Neural*
1095 *Netw.* **19**, 1120–36 (2006).

1096 135. Cavanagh, J. F. *et al.* Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nat.*
1097 *Neurosci.* **14**, 1462–7 (2011).

1098 136. Herz, D. M., Zavala, B. A., Bogacz, R. & Brown, P. Neural Correlates of Decision Thresholds in the Human Subthalamic
1099 Nucleus. *Curr. Biol.* **26**, 916–920 (2016).

1100 137. Montague, P. R., Dayan, P., Person, C. & Sejnowski, T. J. Bee foraging in uncertain environments. *Nature* **377**, 725–728
1101 (1995).

1102 138. Hoerzer, G. M., Legenstein, R. & Maass, W. Emergence of complex computational structures from chaotic neural
1103 networks through reward-modulated hebbian learning. *Cereb. Cortex* **24**, 677–690 (2014).

1104 139. Bailey, C. H., Giustetto, M., Huang, Y. Y., Hawkins, R. D. & Kandel, E. R. Is Heterosynaptic modulation essential for
1105 stabilizing hebbian plasiticity and memory. *Nat. Rev. Neurosci.* **1**, 11–20 (2000).

1106 140. Frank, M. J. Dynamic dopamine modulation in the basal ganglia: A neurocomputational account of cognitive deficits in
1107 medicated and nonmedicated Parkinsonism. *J. Cogn. Neurosci.* **17**, 51–72 (2005).

1108 141. Gerstner, W., Lehmann, M., Liakoni, V., Corneil, D. & Brea, J. Eligibility Traces and Plasticity on Behavioral Time Scales:
1109 Experimental Support of NeoHebbian Three-Factor Learning Rules. *Front. Neural Circuits* **12**, 1–16 (2018).

1110 142. Doya, K. Complementary roles of basal ganglia and cerebellum in learning and motor control. *Curr. Opin. Neurobiol.* **10**,
1111 732–739 (2000).

1112 143. Legenstein, R., Pecevski, D. & Maass, W. A learning theory for reward-modulated spike-timing-dependent plasticity
1113 with application to biofeedback. *PLoS Comput. Biol.* **4**, (2008).

1114 144. Lalive, A. L., Lien, A. D., Roseberry, T. K., Donahue, C. H. & Kreitzer, A. C. Motor thalamus supports striatum-driven
1115 reinforcement. *Elife* **7**, 1–22 (2018).

1116 145. Mcclure, S. M., Berns, G. S. & Montague, P. R. Temporal Prediction Errors in a Passive Learning Task Activate Human
1117 Striatum. **38**, 339–346 (2003).

1118 146. Doherty, J. O. *et al.* Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. **304**, 452–455
1119 (2004).

1120 147. Badre, D. & Frank, M. J. Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: evidence from
1121 fMRI. *Cereb. Cortex* **22**, 527–36 (2012).

1122 148. Lillicrap, T. P., Cownden, D., Tweed, D. B. & Akerman, C. J. Random synaptic feedback weights support error
1123 backpropagation for deep learning. *Nat. Commun.* **7**, 1–10 (2016).

1124 149. O'Reilly, R. C., Munakata, Y., Frank, M. J. & Hazy, T. E. *Computational cognitive neuroscience*. (PediaPress, 2012).

1125 150. Perich, M. G. & Rajan, K. Rethinking brain-wide interactions through multi-region 'network of networks' models. *Curr.*
1126 *Opin. Neurobiol.* **65**, 146–151 (2020).

1127 151. Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C. & Gerstner, W. Inhibitory Plasticity Balances Excitation and Inhibition
1128 in Sensory Pathways and Memory Networks. *Science (80-.).* **334**, 1569–1573 (2011).

1129 152. Fusi, S., Miller, E. K. & Rigotti, M. Why neurons mix: High dimensionality for higher cognition. *Curr. Opin. Neurobiol.* **37**,
1130 66–74 (2016).

1131 153. Badre, D., Bhandari, A., Keglovits, H. & Kikumoto, A. The dimensionality of neural representations for control. *Current*
1132 *Opinion in Behavioral Sciences* **38**, 20–28 (2021).

1133 154. Márton, C. D., Schultz, S. R. & Averbeck, B. B. Learning to select actions shapes recurrent dynamics in the corticostriatal
1134 system. *Neural Networks* **132**, 375–393 (2020).

1135 155. Erlhagen, W. & Schöner, G. Dynamic field theory of movement preparation. *Psychol. Rev.* **109**, 545–572 (2002).

1136 156. Klaes, C., Schneegans, S., Schöner, G. & Gail, A. Sensorimotor Learning Biases Choice Behavior: A Learning Neural Field
1137 Model for Decision Making. *PLoS Comput Biol* **8**, e1002774 (2012).

1138 157. Cisek, P. Integrated neural processes for defining potential actions and deciding between them: a computational
1139 model. *J. Neurosci.* **26**, 9761–70 (2006).

1140 158. Christopoulos, V., Bonaiuto, J. & Andersen, R. a. A Biologically Plausible Computational Theory for Value Integration
1141 and Action Selection in Decisions with Competing Alternatives. *PLOS Comput. Biol.* **11**, e1004104 (2015).

1142 159. Thach, W. T., Goodkin, H. P. & Keating, J. G. The cerebellum and the adaptive coordination of movement. *Annual*
1143 *Review of Neuroscience* **15**, 403–442 (1992).

1144 160. Bastian, A. J. Learning to predict the future: the cerebellum adapts feedforward movement control. *Current Opinion in*
1145 *Neurobiology* **16**, 645–649 (2006).

1146 161. Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T. & Wang, X. J. Task representations in neural networks trained
1147 to perform many cognitive tasks. *Nat. Neurosci.* **22**, 297–306 (2019).

1148 162. Baram, A. B., Muller, T. H., Nili, H., Garvert, M. M. & Behrens, T. E. J. Entorhinal and ventromedial prefrontal cortices

1149 abstract and generalize the structure of reinforcement learning problems. *Neuron* **109**, 713-723.e7 (2021).

1150 163. Whittington, J. C. R. *et al.* The Tolman-Eichenbaum Machine: Unifying Space and Relational Memory through
1151 Generalization in the Hippocampal Formation. *Cell* **183**, 1249-1263.e23 (2020).

1152 164. Collins, A. G. E. & Frank, M. J. Neural signature of hierarchically structured expectations predicts clustering and transfer
1153 of rule sets in reinforcement learning. *Cognition* **152**, 160–169 (2016).

1154 165. Lehnert, L., Littman, M. L. & Frank, M. J. Reward-predictive representations generalize across tasks in reinforcement
1155 learning. *PLoS Comput. Biol.* **16**, 1–27 (2020).

1156 166. Amit, D. J. Neural networks counting chimes. *Proc. Natl. Acad. Sci. U. S. A.* **85**, 2141–2145 (1988).

1157 167. O'Reilly, R. C. & Frank, M. J. Making working memory work: a computational model of learning in the prefrontal cortex
1158 and basal ganglia. *Neural Comput.* **18**, 283–328 (2006).

1159 168. Ratcliff, R. & Frank, M. J. Reinforcement-based decision making in corticostriatal circuits: Mutual constraints by
1160 neurocomputational and diffusion models. *Neural Comput.* **24**, 1186–1229 (2012).

1161 169. Doi, T., Fan, Y., Gold, J. I. & Ding, L. The caudate nucleus contributes causally to decisions that balance reward and
1162 uncertain visual information. *Elife* **9**, 568733 (2020).

1163 170. Mink, J. W. The basal ganglia: Focused selection and inhibition of competing motor programs. *Prog. Neurobiol.* **50**,
1164 381–425 (1996).

1165 171. Alexander, G. E. & Crutcher, M. D. Functional architecture of basal ganglia circuits: neural substrates of parallel
1166 processing. *Trends in Neurosciences* **13**, 266–271 (1990).

1167 172. Franklin, N. T. & Frank, M. J. A cholinergic feedback circuit to regulate striatal population uncertainty and optimize
1168 reinforcement learning. *Elife* **4**, 1–29 (2015).

1169 173. Kim, J. *et al.* Inhibitory Basal Ganglia Inputs Induce Excitatory Motor Signals in the Thalamus. *Neuron* **95**, 1181-1196.e8
1170 (2017).

1171 174. Lo, C. C. & Wang, X. J. Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat.*
1172 *Neurosci.* **9**, 956–963 (2006).

1173 175. Schmitt, L. I. *et al.* Thalamic amplification of cortical connectivity sustains attentional control. *Nature* **545**, 219–223
1174 (2017).

1175 176. Taverna, S., Ilijic, E. & Surmeier, D. J. Recurrent collateral connections of striatal medium spiny neurons are disrupted
1176 in models of Parkinson's disease. *J. Neurosci.* **28**, 5504–5512 (2008).

1177 177. Dobbs, L. K. K. *et al.* Dopamine Regulation of Lateral Inhibition between Striatal Neurons Gates the Stimulant Actions of
1178 Cocaine. *Neuron* **90**, 1100–1113 (2016).

1179 178. Averbeck, B. B., Sohn, J. W. & Lee, D. Activity in prefrontal cortex during dynamic selection of action sequences. *Nat.*

1180 *Neurosci.* **9**, 276–282 (2006).

1181 179. Berdyyeva, T. K. & Olson, C. R. Monkey Supplementary Eye Field Neurons Signal the Ordinal Position of Both Actions
1182 and Objects. *J. Neurosci.* **29**, 591–599 (2009).

1183 180. Isoda, M. & Tanji, J. Contrasting Neuronal Activity in the Supplementary and Frontal Eye Fields during Temporal
1184 Organization of Multiple Saccades. *J. Neurophysiol.* **90**, 3054–3065 (2003).

1185 181. Isoda, M. & Tanji, J. Participation of the primate presupplementary motor area in sequencing multiple saccades. *J.*
1186 *Neurophysiol.* **92**, 653–659 (2004).

1187 182. Shima, K. & Tanji, J. Neuronal activity in the supplementary and presupplementary motor areas for temporal
1188 organization of multiple movements. *J. Neurophysiol.* **84**, 2148–60 (2000).

1189 183. Clower, W. T. & Alexander, G. E. Movement sequence-related activity reflecting numerical order of components in
1190 supplementary and presupplementary motor areas. *J. Neurophysiol.* **80**, 1562–1566 (1998).

1191 184. Salinas, E. Rank-Order-Selective Neurons Form a Temporal Basis Set for the Generation of Motor Sequences. *J.*
1192 *Neurosci.* **29**, 4369–4380 (2009).

1193 185. Schwartze, M., Keller, P. E., Patel, A. D. & Kotz, S. A. The impact of basal ganglia lesions on sensorimotor
1194 synchronization, spontaneous motor tempo, and the detection of tempo changes. *Behav. Brain Res.* **216**, 685–691
1195 (2011).

1196 186. Gershman, S. J., Moustafa, A. A. & Ludvig, E. A. Time representation in reinforcement learning models of the basal
1197 ganglia. *Front. Comput. Neurosci.* **7**, 1–8 (2014).

1198 187. Jones, C. R. G. & Jahanshahi, M. Contributions of the Basal Ganglia to Temporal Processing: Evidence from Parkinson's
1199 Disease. *Timing Time Percept.* **2**, 87–127 (2014).

1200 188. Thura, D. & Cisek, P. The Basal Ganglia Do Not Select Reach Targets but Control the Urgency of Commitment. *Neuron*
1201 **95**, 1160-1170.e5 (2017).

1202 189. Paton, J. J. & Buonomano, D. V. The Neural Basis of Timing: Distributed Mechanisms for Diverse Functions. *Neuron* **98**,
1203 687–705 (2018).

1204 190. Gurney, K., Prescott, T. J. & Redgrave, P. A computational model of action selection in the basal ganglia. I. A new
1205 functional anatomy. *Biol. Cybern.* **84**, 401–410 (2001).

1206 191. Alexander, G. E., DeLong, M. R. & Strick, P. L. Parallel organization of functionally segregated circuits linking basal
1207 ganglia and cortex. *Annu. Rev. Neurosci.* **VOL. 9**, 357–381 (1986).

1208 192. Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508**, 207–214 (2014).

1209 193. Hintiryan, H. *et al.* The mouse cortico-striatal projectome. *Nat. Neurosci.* **19**, 1100–1114 (2016).

1210 194. Hunnicutt, B. J. *et al.* A comprehensive excitatory input map of the striatum reveals novel functional organization. *Elife*

1211 **5**, 1–32 (2016).

1212 195. Peters, A. J., Fabre, J. M. J., Steinmetz, N. A., Harris, K. D. & Carandini, M. Striatal activity topographically reflects
1213 cortical activity. *Nature* **591**, (2021).

1214 196. Gerardin, E. *et al.* Foot, hand, face and eye representation in the human striatum. *Cereb. Cortex* **13**, 162–169 (2003).

1215 197. McHaffie, J. G., Stanford, T. R., Stein, B. E., Coizet, V. & Redgrave, P. Subcortical loops through the basal ganglia. *Trends*
1216 *Neurosci.* **28**, 401–7 (2005).

1217 198. Znamenskiy, P. & Zador, A. M. Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination.
1218 *Nature* **497**, 482–485 (2013).

1219 199. Friedman, A. *et al.* A corticostriatal path targeting striosomes controls decision-making under conflict. *Cell* **161**, 1320–
1220 1333 (2015).

1221 200. Gremel, C. M. *et al.* Endocannabinoid Modulation of Orbitostriatal Circuits Gates Habit Formation. *Neuron* **90**, 1312–
1222 1324 (2016).

1223 201. Hooks, B. M. *et al.* Topographic precision in sensory and motor corticostriatal projections varies across cell type and
1224 cortical area. *Nat. Commun.* **9**, (2018).

1225 202. Lee, J., Wang, W. & Sabatini, B. L. Anatomically segregated basal ganglia pathways allow parallel behavioral
1226 modulation. *Nat. Neurosci.* (2020). doi:10.1038/s41593-020-00712-5

1227

1228