A decision-space model explains context-specific decision-making

Alexander Friedman

afriedman@utep.edu

The University of Texas at El Paso https://orcid.org/0009-0006-0882-5324

Dirk Beck

The University of Texas at El Paso

Cory Heaton

The University of Texas at El Paso

Luis Davlia

The University of Texas at El Paso

Lara Rakocevic

The University of Texas at El Paso

Sabrina Drammis

MIT

Danil Tyulmankov

University of Southern California

Paulina Vara

The University of Texas at El Paso

Atanu Giri

The University of Texas at El Paso

Shreeya Umashankar Beck

The University of Texas at El Paso

Qingyang Zhang

The University of Texas at El Paso

Michael Pokojovy

Old Dominion University

Kenichiro Negishi

National Institute on Drug Abuse

Serina Batson

The University of Texas at El Paso

Alexis Salcido

The University of Texas at El Paso

Neftali Reyes

The University of Texas at El Paso
Andrea Macias
The University of Texas at El Paso
Raquel Ibanez-Alcala
The University of Texas at El Paso
Safa Houssain
The University of Texas at El Paso
Graham Waller
The University of Texas at El Paso
Travis Moschak
The University of Texas at El Paso
Ki Goosens
United States https://orcid.org/0000-0002-5246-2261

Article

Keywords:

Posted Date: December 3rd, 2024

DOI: https://doi.org/10.21203/rs.3.rs-5499511/v1

Additional Declarations: There is **NO** Competing Interest.

A decision-space model explains context-specific decision-making

Authors

Beck, Dirk W.[2]; Heaton, Cory N.[1]; Davila, Luis D.[2#]; Rakocevic, Lara I.[2#]; Drammis, Sabrina M.[3#]; Tyulmankov, Danil[5#]; Vara, Paulina[1]; Giri, Atanu[2]; Umashankar Beck, Shreeya[1]; Zhang, Qingyang[1]; Pokojovy, Michael[6]; Negishi, Kenichiro[4]; Batson, Serina A.[1]; Salcido, Alexis A.[1]; Reyes, Neftali F.[1]; Macias, Andrea Y.[1]; Ibanez-Alcala, Raquel J.[1]; Hossain, Safa B.[1]; Waller, Graham L.; Moschak, Travis M.[1]; Goosens, Ki A.*[7]; Friedman, Alexander*[1,2,8]

1. Department of Biological Sciences, University of Texas at El Paso, El Paso, TX, USA
2. Computational Science Program, University of Texas at El Paso, El Paso, TX, USA
3. Artificial Intelligence Laboratory, Department of Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA
4. National Institute on Drug Abuse, Baltimore, MD, USA
5. Ming Hsieh Department of Electrical and Computer Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, CA
6. Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA, USA
7. Department of Psychiatry; Center for Translational Medicine and Pharmacology; Friedman Brain Institute; Icahn School of Medicine at Mount Sinai, NY, USA
8. Lead author

# Denotes 3rd authors who have contributed equally to the manuscript

*. Corresponding authors, please address ki.goosens@mssm.edu and afriedman@utep.edu.

Abstract

Optimal decision-making requires consideration of internal and external contexts. Biased decision-making is a transdiagnostic symptom of neuropsychiatric disorders. We created a computational model demonstrating how the striosome compartment of the striatum constructs a context-dependent mathematical space for decision-making computations, and how the matrix compartment uses this space to define action value. The model explains multiple experimental results and unifies other theories like reward prediction error, roles of the direct versus indirect pathways, and roles of the striosome versus matrix, under one framework. We also found, through new analyses, that striosome and matrix neurons increase their synchrony during difficult tasks, caused by a necessary increase in dimensionality of the space. The model makes testable predictions about individual differences in disorder susceptibility, decision-making symptoms shared among neuropsychiatric disorders, and differences in neuropsychiatric disorder symptom presentation. The model provides new evidence for the central role that striosomes play in neuroeconomic and disorder-affected decision-making.

Main

Decision-making is altered in neuropsychiatric disorders affecting the basal ganglia[1]. A range of experimental evidence links, in particular, balances between the compartments of the striatum and connected brain regions to decision-making function and dysfunction[2–6]. Understanding these intricate interactions will be crucial for designing next-generation treatments.

Striatal neurons can be categorized via neurochemistry and connectivity into groups, including the striosome and matrix compartments[4,7]. Striosomal spiny projection neurons (sSPNs) make up ~10-15% of the striatum and matrix spiny projection neurons (mSPNs) another ~85-90%[8,9] (for acronyms, see **Table 1**). New technologies, including recording and targeting methods and genetically engineered mice, have enabled important new discoveries about differential roles for striosomes and matrix in decision-making[3,10–19], including in disorders[2,4]. Further, both sSPNs and mSPNs belong to either the direct pathway (dsSPNs and dmSPNs), identified by D1 receptor expression, or the indirect pathway (dsSPNs and isSPNs) identified by D2 receptor expression[4,20]. Notably, dsSPNs, to a greater extent than dmSPNs, project in turn to regions which influence midbrain dopamine release via a subcircuit that is conserved across species[8,16,21–24] (**Fig. 1a, Table 2**). Thus, dsSPNs, isSPNs, dmSPNs, and imSPNs appear to have different physiological roles[16,20], raising the possibility they each play a distinct functional role during decision-making.

While a range of modeling works explore the roles of direct versus indirect pathway function[25], they largely omit the important dimension of striosome versus matrix. The omission of striosomes from models of decision-making or basal ganglia function hinders the interpretation of important features of the striatum because it prevents an accurate depiction of striatal-daSNC interplay, which is primarily striosomal[4,24,26] (**Table 2**). Further, disorders that differentially affect the direct versus indirect pathways have been found to also affect striosomes versus matrix differently[2], suggesting that attention to all four compartments is necessary for an accurate understanding of disorder-affected decision-making. To close this gap, we formed a model that accounts for striosome versus matrix subdivisions, including the selective modulation of midbrain dopamine by striosomes (**Fig. 1, Extended Data Fig. 1**). From our physiological model arises the concept of a "decision-dimension," our term for an axis along which the modeled circuit encodes information (for terminology, see **Table 1**). During a decision, decision-dimensions that are important based on the context are selected, forming a mathematical "decision-space." We present evidence, via our analysis of the neural recordings and our models of findings from the experimental literature, for the core tenets of our model: that subpopulations of SPNs encode information along decision-dimensions, that a decision-space is formed, and that the decision-space adapts based on context (**Figs. 2,3, Extended Data Figs. 2-5**). Then we demonstrate the power of the model to explain a range of physiological and behavioral phenomena, including RPE and the roles of the indirect/direct pathway (**Fig. 4, Extended Data Fig. 6**). Finally, we speculate how the model might explain behavioral phenomena observed in psychiatry (**Figs. 5,6, Extended Data Figs. 7,8**) and suggest future experiments (**Extended Data Fig. 9**).

Results

Model description.

Our model describes how physiological interactions between elements of a striosome-centered circuit inform decision-making. (For extended reasoning behind our choice of circuit elements, see **Supplementary Note 1**.) **Striatum-projecting cortical neurons** that encode

mixed information serve as the input to the modeled circuit. We assume for the purpose of our model that the information passed from the cortex to each striatal compartment is roughly similar (in actuality, there is much overlap with some differences, see **Table 2** and **Supplementary Note 2**). The cortical neurons synapse on subpopulations of proximate SPNs which have been found to each encode distinct information[27]. During this process, fast spiking interneurons (**FSIs**) perform a normalization operation (**Extended Data Fig. 1a**). Mathematically, we represent this as a matrix $\mathbf{W}_P$ dictating cortex-SPN connection weights for each pathway P (direct or indirect) mapping cortical activity $\mathbf{x}_P$ to the coordinate space of sSPNs, where it undergoes divisive normalization by FSI activity c_P and a shift in activity b_{sSPN} to form sSPN activity $\mathbf{s}_{\text{sSPN},P}$:

(1) $$\mathbf{s}_{\text{sSPN},P} = \frac{1}{c_P}\mathbf{W}_P{}^{\mathrm{T}}\mathbf{x}_P + b_{\text{sSPN}}$$

Information from sSPNs is then passed to dopaminergic neurons of the substantia nigra compacta (**daSNC**). daSNC neurons, like sSPNs, have been shown to be organized topologically into subpopulations that each encode distinct information[28]. Experimental evidence suggests that signals are passed from sSPN subpopulations to daSNC neurons in three ways (see **Table 2**): **A) dsSPN→daSNC**. Subpopulations of dsSPNs inhibit daSNC subpopulations directly via dendrite bouquets[26]. Thus, in our model, each dsSPN subpopulation inhibits a corresponding daSNC subpopulation. **B) isSPN→GPe→daSNC**. isSPNs send signals to daSNC neurons via GPe[16], which inhibit daSNC subpopulations. Thus, in our model, each isSPN subpopulation disinhibits a corresponding daSNC subpopulation. **C) sSPN→GPi→LHb→RMTg→daSNC.** GPi integrates signals from many sSPN subpopulations through synapses that release both GABA and glutamate[29], and the LHb, when activated by GPi, powerfully inhibits multiple of the dopaminergic subpopulations via RMTg[30,31]. So, in our model, shifts in this pathway lead to a shift across all daSNC subpopulations. Mathematically, we represent the three circuits as daSNC combining activity from sSPNs $\mathbf{s}_{\text{sSPN},P}$ (with connection weights $w_{\text{sSPN}\to\text{daSNC},i,P}$ corresponding to each decision-dimension i and pathway P) with RMTg activity and an additive shift $z_{\text{daSNC},i,P}$:

(2) $$\text{daSNC}_{i,P} = \frac{1}{1+\exp\left(w_{\text{sSPN}\to\text{daSNC},i,P}\cdot \mathbf{s}_{\text{sSPN},i,P} + \text{RMTg} - z_{\text{daSNC},i,P}\right)}$$

In addition to sSPNs, there are subpopulations of **mSPNs**, termed matrisomes[4], that densely surround sSPN subpopulations. In our model, we hypothesize that these sSPN and mSPN subpopulations communicate with one another via dopamine release from the corresponding daSNC subpopulation (there are other sSPN→mSPN connections which we do not model that play more local roles, see **Supplementary Note 3**). There are multiple groups of these functionally connected sSPNs, daSNCs, and mSPNs. In the model, when a daSNC subpopulation is active, dopamine is released to the corresponding sSPN and mSPN subpopulations, resulting in enhanced or inhibited mSPN reception of cortical signal among the subpopulations, as shown in experimental work[32]. Mathematically, mSPN activity is defined similarly to sSPN activity, but for a diagonal matrix $\mathbf{S}_P$ corresponding to the dopamine release that probabilistically defines the decision-space, with $\mathrm{P}\left(\mathbf{S}_{\mathbf{P},\mathrm{ii}} = 1\right) = \text{daSNC}_{i,\mathrm{P}}$:

(3) $$\mathbf{s}_{\text{mSPN},P} = \frac{1}{c_P}\,\mathbf{S}_P\mathbf{W}_P{}^{\mathrm{T}}\mathbf{x}_P$$

mSPNs have been found to be primarily involved in motor functions, projecting to the GPi, SNr, and then to brainstem motor programs[33]. SPN activity (which is predominantly mSPN) has been shown to contribute to action selection and initiation[34]. The direct pathway has generally been implicated in promoting actions and the indirect pathway in preventing actions[35]. So, in our model, the output of the circuit is the definition of action values (the value of performing various actions, encoded by the direct pathway) and inaction values (the value of refraining from those actions, encoded by the indirect pathway) by mSPN signals on route to downstream regions. Mathematically, values $v_{j,P}$ (for each action/inaction j, pathway P) are defined based on mSPN activity, internal coefficients $\beta_{j,P}$, and priors $\alpha_{j,P}$ (which are set to arbitrary values that are constant across analyses, see **Common parameters, Methods**):

(4) $$v_{j,P} = \frac{1}{1 + \exp\left(-\beta_{j,P}\, s_{\mathrm{mSPN},P} - \alpha_{j,P} \right)},$$

Based on these values, actions are either performed or refrained from over time. We model this using a Merton process model where the first process to hit a threshold is enacted (direct pathway) or refrained from (indirect pathway). See **Defining choice, Methods**.

Thus, our model is constructed based on the anatomy and physiology of the striosome-centered circuit. The physiological description also produces a simple and convenient geometric interpretation. If we let each SPN and daSNC subpopulation encode the principal components of cortical activity, such as could be learned via a modified Oja's rule[36], then the columns of $\mathbf{W}_P$ become orthogonal. So, each SPN subpopulation can be thought of as encoding information along an axis of Euclidean space. We term these axes "decision-dimensions" and have evidence that they correspond to constructs such as reward, cost, or novelty (discussed in more detail below). We suggest that when dopamine is released to SPNs, selectively enhancing reception of cortical signal, decision-dimensions are effectively prioritized. Therefore, cortex, striosomes, and dopamine work together to form a "decision-space", only focusing on decision-dimensions that are necessary based on context (**Figs. 1b,c**). In this light, sSPN and daSNC, via pairs of connected subpopulations that process information in parallel, serve the functional role of selecting *which* decision-dimensions should receive high priority. In particular, dsSPNs determine which decision-dimensions to use in the direct pathway, and isSPNs which decision-dimensions to use in the indirect pathway. On the other hand, GPi, LHb, and RMTg, by prioritizing or deprioritizing all daSNC subpopulations together, determine *how many* decision-dimensions should be prioritized (**Fig. 1d, Extended Data Figs. 1b-i**).

Importantly, the advantage of this formation is not only conceptual, but practical for linking physiology to decision-making (**Figs. 1e-g, Extended Data Figs. 1j-p**). Distinct sSPN and/or daSNC subpopulations have been found to encode, for instance, reward[12,14,15,28], cost[12,14,15,28], or novelty[37]. Thus, we might imagine that decision-dimensions could correspond loosely to reward level, cost level, or novelty level. If this is the case, a logical prediction of the model arises: we would expect a low-dimensional decision-space to be formed during a simple choice (e.g. between two rewards) and a high-dimensional decision-space to be formed during a more difficult choice (e.g. between offers which each have benefits and costs that must be weighed in order to solve the problem). This hypothesis, if proven, would allow us to infer the decision-spaces of behaving rodents or humans simply by regressing sSPN activity on experimental parameters (for example, temperature or music volume), as we demonstrate using synthetic data (**Extended Data**

Figs. 1q,r). For example, a significant correlation between sSPN activity and novelty level would indicate the existence of decision-dimension that corresponds roughly to novelty (**Inferring decision-space from SPN activity and choice, Methods**). We sought to determine if this hypothesis is supported by experimental physiological data collected during decision-making.

Support for the model: context-dependent sSPN physiology matches model predictions.

We began by asking whether the results of the physiological sSPN literature support a link between the decision-space (which our model postulates is driven by sSPN activity) and task difficulty (inferred from experimental inputs, for instance a simple task with reward only versus a difficult task with conflicting rewards and costs). We began with the experimental literature on sSPNs. In one experiment[11], sSPNs were optogenetically stimulated (or inhibited) during a rodent conflict decision-making task. Per our model, this should cause inhibition (or disinhibition) of daSNC neurons, leading to reduced (or enhanced) dopamine release to SPNs, producing a lower- or (higher-) dimensional decision-space. Indeed, the stimulation led to choices indicative of decision-making using few informational dimensions, while inhibition led to choices indicative of decision-making using multiple informational dimensions (**Figs. 2a,b, Extended Data Fig. 2a, Effect of sSPN activity on decision-space** and **Effect of decision-space on choice, Methods**). A second experiment[11] tested, reversely, the activity of striosomes during simple versus difficult tasks. As the model expects, striosomal activity during simple tasks resembled the levels observed during optogenetic excitation, whereas striosomal activity during difficult tasks resembled the levels observed during optogenetic inhibition (**Fig. 2c, Extended Data Figs. 2b-d**). In another study, striosome activity was lower among rodents that learned a difficult reversal learning task than among rodents that did not[10]. Thus, there appears to be a relationship between sSPN physiology and task difficulty in the direction expected by our model.

Support for the model: context-dependent sSPN-mSPN synchrony matches model predictions.

Importantly, the result above does not distinguish our model from the alternative, simpler explanation that sSPNs might collectively encode task difficulty. We could term such a model a "conflict model," where sSPN activity tracks the overall conflict present in a task (see **Table 3**). To determine whether it was changes to sSPN subpopulations driving the overall change in sSPN activity (i.e. decision-space), rather than a general effect, we analyzed the paired activities of sSPNs during simple versus difficult tasks using the Corticostriosomal Circuit Stress Experimental database[3]. We hypothesized that a greater number of sSPNs and mSPNs would be functionally connected during difficult tasks. This prediction arises from the circuit connectivity in our model, where sSPNs are functionally connected to mSPNs via daSNC, and each additional prioritized subpopulation causes more sSPN→daSNC→mSPN modulation (**Fig. 2d**). This could be observed, for instance, as an increase in striosome and matrix correlation as decision-space increases. This hypothesis is also inspired by the observation that striosome and matrix activity has been found to roughly track one another over time in a difficult task[13].

To test this, we analyzed the synchrony between striosome neurons and matrix neurons during simple decisions (that forced choice between either two rewards or two costs) and during difficult decisions (that forced choice between offers that contained both rewards and costs together). We measured synchrony as the cross-correlation between sSPN activity and mSPN activity over the period of the task. To control for possible physiological differences across the phases of the rats' movements, we also developed a custom Granger causality-based tool. Per both metrics, synchrony was significantly higher in the difficult task. Further, synchrony scaled with the difficulty with which the rats treated the task, as measured based on deliberation time (**Fig. 2e, Extended Data Figs. 2e-n, Connected SPNs through cross-correlation** and

Connected SPNs through Granger causality, Methods). Thus, rather than sSPNs encoding conflict in their general level of activity, there appears to be an important relationship between sSPN and mSPN subpopulations during high-conflict tasks. Our analysis does not confirm that sSPNs have a causal effect on mSPNs. However, if this is assumed, the evidence suggests enhanced modulation of mSPN subpopulations by sSPN subpopulations during the difficult tasks, that is, a formation of a higher-dimensional decision-space.

Support for the model: Dimensionality reduction from the cortex to SPNs.

Notably, the synchrony analysis above demonstrates the use of subpopulations during context-dependent decision-making, but it does not test whether those subpopulations correspond to decision-dimensions. There is, however, evidence that SPNs encode what we term decision-dimensions. Experimental work has demonstrated that distinct SPN subpopulations encode different information, and that these subpopulations persist across days[27]. Further, our analysis suggests that dimensionality reduction occurs from the cortex to SPNs, as it does in our model during the mapping of information to a basis of decision-dimensions. Cortical neurons had the most coordinated activities over time (measured as effective correlation[38]), then FSIs, followed by sSPNs and mSPNs **(Fig. 2f, Analyzing neural dimensionality reduction, Methods)**. This would suggest a higher-dimensional representation in cortex, where neurons encode similar information over time, than in the downstream regions, where information is compactly encoded in neurons that behave differently over time. Thus, it seems that cortex→SPN dimensionality reduction occurred during the tasks, like in our model the mapping from high-dimensional cortical information to a basis of SPN decision-dimensions.

Support for the model: Strong alignment to the experimental literature on SPNs compared to alternative models.

The analysis to this point is based on a selection of the experimental literature relating sSPN activity to mSPN activity and choice. We wished to test the model more broadly using a range of experimental results. To this end, we devised five tests that link sSPN activity to choice, each verifiable with the experimental literature, that might support or reject the decision-space model (**Extended Data Figs. 2o-x, Table 3**). As benchmarks, we also constructed, from roles commonly assigned to sSPNs, four alternative models in which sSPNs encode 1) conflict, 2) subjective value, 3) prediction error, or 4) actions. We found that while the alternative models each can be used to interpret a subset of the experimental evidence, only the decision-space model aligned to the breadth of it (**Table 4**). A selection of experimental studies on GPi, LHb, and daSNC also align with the decision-space model (see **Tables 5-8**), thus offering a new lens through which to interpret their functions.

Inability to form a high-dimensional decision-space in disorders.

We next applied our model to the experimental literature on neuropsychiatric disorders, wondering if it could offer insight into disorders associated with sSPN changes. Interestingly, in an experimental study on chronic stress[3], sSPNs were hyperactive compared to controls during the most difficult task but less affected during the simpler tasks (**Fig. 3a**). Meanwhile, the rodents were less adherent to reward level only in the difficult task, suggesting dysfunction in processing reward and cost together (**Figs. 3b,c, Effect of decision-space on choice, Methods**). Thus, we wondered if post-stress sSPN hyperactivity could cause altered choices due to a reduction in the dimensionality of the decision-space, similar to the model in **Fig. 2b** where optogenetic excitation reduced the decision-space in controls.

Our analysis supports this. Animals made decisions in the cost-benefit conflict task more quickly after stress, as if the task were less difficult (**Extended Data Figs. 3a-d, Defining decision difficulty by task, Methods**). Meanwhile, after stress, choices involving both reward and cost no longer had more functionally connected sSPNs and mSPNs than the simple tasks, suggesting a change to the decision-space as well as a general shift in activity. In fact, synchrony was similar across tasks and to the simple tasks for controls (**Extended Data Figs. 3e-i**). Thus, after stress, the rodents both showed both neural signatures aligned with a low-dimensional decision-space and choices aligned with processing of information in a low-dimensional way.

The inability to form a high-dimensional decision-space can also explain a counterintuitive finding that stress causes rodents to prefer a reward-cost combination over a reward presented without cost (**Extended Data Fig. 3j, Changes to choice after adding cost to a reward offer, Methods**). In classic economic theory, the addition of cost to a reward typically makes a good less attractive[39] and thus our observation cannot be readily explained. In contrast, the decision-space model offers a simple explanation: cost can increase offer attractiveness in instances where cost level causes a transition from a default low-dimensional decision-space to a higher-dimensional decision-space, as we hypothesized is the case after stress in **Figs. 3b,c**. In these cases, the rules encoded by mSPNs can assign a higher value to accepting versus avoiding an offer when reward and cost are considered rather than only reward (**Figs. 3d,e**). In other cases, decisions are predicted to resemble those predicted by classic theory, for example in cases where either a one-dimensional (as in **Fig. 3d**) or two-dimensional (**Fig. 3e**) decision-space is used across cost levels.

Cortex→FSI connectivity is impacted by chronic stress[3] (**Extended Data Figs. 3k-r, Analyzed cortex-FSI connectivity** and **Modeled cortex-FSI connectivity after stress, Methods**), leading to hyperactive sSPNs. Our model suggests that this causes the formation of a lower-dimensional decision-space (**Figs. 3f,g**), which would lead to lower variance and higher mean of SPN subpopulations (**Extended Data Fig. 4**). Notably, the cortex→FSI connection is also impacted in Huntington's disease and aged rodents[10], raising the possibility that an inability to form a high-dimensional decision-space during difficult decisions is a feature of multiple health conditions. Supporting this hypothesis, a model where Huntington's disease and aged subjects use a lower-dimensional decision-space produces action values that follow the trend of experimental choice (**Fig. 3h, Effect of decision-space on choice, Methods**).

Interestingly, our model expects low dimensional decision-space to be beneficial in disorder conditions. A feature of chronic stress[3] and schizophrenia[40] is reduced cortical signal-to-noise ratio (SNR) and disrupted cortical signaling. In these conditions, a low dimensional decision-space is theoretically optimal because only the highest-priority decision-dimensions carry enough signal to outweigh the drawback of noise (**Extended Data Fig. 5, Effect of cortical SNR on choice, Methods**). Our analysis raises the possibility that FSIs help steer the circuit towards a helpful decision-space in disorders.

Functional roles of dsSPNs, isSPNs, dmSPNs, and imSPNs.

sSPNs and mSPNs have been found to be distributed between the direct and indirect pathways of the striatum[24,41,42], with dopamine release differently affecting sSPNs versus mSPNs and also dSPNs versus iSPNs[20,32,43]. Various neuropsychiatric disorders are associated with disturbed sSPN versus mSPN[2] and direct versus indirect pathway balances[44]. Thus, the compartments likely play distinct functional roles in decision-making, including in disorders.

Because dsSPNs connect to dmSPNs through daSNC in our model, and isSPNs to imSPNs through GPe and daSNC, two decision-spaces are formed in parallel, one related to each pathway (**Fig. 4a**). Thus, based on the functional roles we assign the pathways, the circuit uses a direct pathway decision-space to determine whether to perform an action and an indirect pathway decision-space to determine whether to refrain from it. dsSPNs influence the direct pathway decision-space while isSPNs influence the indirect pathway decision-space, and dmSPNs promote actions while imSPNs discourage actions (**Figs. 4a-e, Modeling time-variant input, Methods**). The circuit uses these compartment-specific mechanisms to calculate which actions *should* be performed with one set of decision-dimensions and calculate which actions *should not* be performed with another. The answers to these questions might overlap. For instance, the direct-pathway and indirect-pathway space should provide divergent answers to the value of consuming cocaine based on the dimensions they prioritized; the former, focusing on reward, might assign it great value while the latter, focusing on cost, might assign great value to not consuming it. When balances between the direct versus indirect pathways and striosome versus matrix changes, this calculus changes. For example, it has been found that dopamine release is enhanced to sSPNs versus mSPNs after cocaine administration[45] and simultaneously dSPNs are enhanced in the short-term and iSPNs over longer-term scales of time[46]. This might lead to a high-dimensional direct-pathway decision-space but pruning of ordinarily important decision-dimensions from the indirect pathway decision-space, producing a reduction of nuance in determining when to avoid actions and heightened impulsivity.

Indeed, the model's interpretation of the two parallel decision-spaces offers an intuitive explanation for a range of experimental observations on the direct versus indirect pathway. For example, our model replicates the experimental observations that increased dopamine leads to riskier and quicker decisions and preference for nearby offers, in time or physical proximity, compared to distant offers (**Extended Data Fig. 6, Table 8, Effect of dopamine on action/inaction values** and **Effect of decision-dimensions on choice, Methods**).

Prediction error encoding is an emergent property of the model.

A range of experimental studies have shown that SPN activities track prediction errors[12,47]. This observation has led to hypotheses that SPNs encode a function related to prediction errors in a reinforcement learning framework[12]. The decision-space model offers a different explanation. In our model, the weights from cortical neurons to SPNs naturally separate cortical information by their associations. For instance, if a bell tends to sound when a subject drinks chocolate milk, both stimuli, even if they arrive from different cortical sources, will likely be mapped to the same reward-related decision-dimension as synapses adjust per Oja's rule. Less reliable cues are expected to develop mappings with smaller weights. Therefore, the activities of reward-related SPNs may rise when cues predicting rewards appear and fall when cues predicting less reward appear. A sudden change to reward information, for instance a predictive cue, should thus lead to a sudden change in the activity of an SPN subpopulation related to reward, and this change in activity should resemble a prediction error (**Figs. 4f-h**). Thus, in contrast with the more traditional interpretation where SPNs internally encode a temporal difference value function, the decision-space model suggests that the mapping of cortical activity to the basis of striatal decision-dimensions is sufficient to track prediction error in many cases, without additional computational work performed by the SPNs.

Both interpretations can be used to explain much of the experimental evidence, although the interpretation of the decision-space model may more closely align to recent experimental data (**Tables 3,4**). Our model also may provide a functional rational for the observation that separate SPNs encode data along different informational axes[12,14]. In fact, our model expects more of these

axes to be uncovered by future work. We might expect, for instance, a cue predictive of the novelty of an object to produce an immediate change in activity of an SPN subpopulation related to novelty.

How might the circuit respond in cases where new information diverges sharply from expectations? Our model predicts that in these cases, sSPNs will signal to daSNC that a decision-dimension should be reprioritized, effectively adding or removing the dimension from the decision-space. Interestingly, the circuit has an inherent physiological mechanism to quickly transition away from a decision-space that is no longer optimal. Experimental evidence has identified rebounds in daSNC activity[21] and striatal dopamine release[18] after sSPN optogenetic stimulation. Our model suggests that these observations are part of a system by which the circuit can rapidly de-prioritize a decision-dimension after a negative prediction error (e.g. less reward than expected). Thus, the circuit is able to quickly shift to a more helpful decision-space (**Fig. 4i**).

Discussion

We found evidence, through the experimental literature and our analysis of neural recordings, to support our hypothesis that modeled physiological patterns in SPN activity (the decision-space) can be used to predict patterns in decision-making, and vice versa. This supports our model of the roles of striosomes and matrix neurons of the direct and indirect pathways in context-dependent decision-making.

Due to the circuit's important role in decision-making processes including in neuropsychiatric disorders, our model provides a framework with which to study decision-making phenomena commonly observed in psychiatry. An important prediction of our model is variance in context-dependent decisions, between individuals and over time (**Figs. 5a-e, Extended Data Fig. 7a-e, Effects of sSPN, LHb, and daSNC activity on choice profiles, Methods**). Individual differences in decision-making as a function of disorders, as seen in the experimental literature[48–50], could arise in cases where there are slight differences in activity of the circuit we model, leading to similar decision-making phenotypes only when a similar decision-space is formed. Daily variance in decision-making, a common observation in pscyhology[51], could arise from daily variance in circuit activity, causing daily variance in the decision-spaces formed most often. Further, differences in circuit activity may explain the established inter-individual differences in the severity of psychiatric disorder symptoms observed during decision-making[52,53]. Individual differences in disorder susceptibility could arise from reliance upon or avoidance of a decision-space that leads to extreme decision-making tendencies (e.g. extremely action-heavy, extremely risk-averse) when combined with abnormal action value rules in mSPNs (**Fig. 5f and Extended Data Fig. 7f**).

Further, our model serves as a framework for forming hypotheses about changes to the circuit across days and weeks, including during neuropsychiatric disorder progression. Our model expects the circuit to adapt between trials as it adjusts to more frequently form a preferred decision-space (**Figs. 6a-f, Effect of initial circuit activity on future trials, Methods**). So, vulnerability or resilience to disorders can be framed as an adaptation that is favorable (e.g., to adeptly form a high-dimensional decision-space) or maladaptive (e.g., to only form a low-dimensional decision-space, regardless of decision context). Depending on its initial activity, a modeled circuit can adapt to reach very different activity, leading to disposition to either a high- or low-dimensional decision-space (**Fig. 6g and Extended Data Fig. 8a**). Thus, differences in the circuit before exposure to a traumatic event, for instance, may explain why two subjects that encounter the same traumatic event do not always develop the aberrant decision-making

symptoms of PTSD[54]. It may also shed light on the neural processes underlying incubation of fear[55] and incubation of craving[56], where disorders progress over the span of weeks or months, even when the traumatic event or addictive substance does not reappear (**Extended Data Figs. 8b-d, Effect of altered advantage score on future trials, Methods**). By representing the role of SPNs in a compartment-specific way, our model facilitates understanding of disorders that affect striosomes and matrix differentially.

Our model adds detail to the predictions made by a range of other models of basal ganglia function. For example, a range of other models consider how the direct and indirect pathways of the basal ganglia interact to moderate action selection[25,57]. Our model offers a simple explanation for several experimental phenomena that are used to fit other models by differentiating the roles of striosomes versus matrix (**Extended Fig. 6**). In particular, direct/indirect pathway models more commonly fit data from simple reward and/or cost tasks rather than highly difficult tasks, and our model allows for scaling to difficult tasks with many informational dimensions. Another class of model examines how RPE signals facilitate adaptability[58–60]. Our model does not preclude the possibility that SPNs and/or dopamine may encode an algorithm similar to those used in reinforcement learning. However, our model expects that such an algorithm would function within the framework of the decision-space, perhaps by encoding a separate reinforcement learning value function along each decision-dimension, instead of a traditional model that learns only the discounted sum of future rewards. Exploring this possibility would add depth to current studies of reinforcement learning in the basal ganglia. Our model can additionally be used to expand on other models that explore the role of basal ganglia pathways in performing dimensionality reduction[61] or responding to events as sequences that influence each successive action[62,63]. In each of these cases, our model adds detail on the functional process by which the basal ganglia prepares for a decision based on context. It also clarifies the mechanisms by which the compartments of the striatum process information, led by striosomal influence on dopamine.

Our model carries several limitations, including its limited focus on a dorsomedial striosomal circuit and certain physiological assumptions (**Supplemental Notes 1-4**). We limit our focus to a specific circuit that has been implicated in decision-making, rather than attempt a unifying theory of basal ganglia function or decision-making encoding across the brain (**Supplemental Note 1**). While we demonstrate alignment to the existing experimental evidence in **Figs. 2 and 3 and Tables 3-8,** future experiments (outlined in **Extended Data Fig. 9**) will be required to confirm the several assumptions we make. Despite these limitations, our model has demonstrated success in relating neural activity to decision-making across a range of behavioral tasks and has the power to explain a range of phenomena, from neural processes to psychiatric observations.

Acknowledgements

We greatly appreciate G. Schoenbaum, and Y. Shaham for their constructive criticism and discussions during model development. This project was supported by the NSF/CAREER (#2235858), NIH/NIDA (#R01DA058653), U-RISE T34GM145529 and G-RISE T32GM144919.

Author Contributions

conceptualization: A.F., K.A.G.; data curation: D.W.B., L.D.D., L.R.I., S.B.D., A.G. Q.Z.; formal analysis: D.W.B., L.R.I., L.D.D., C.N.H., A.G., Q.Z., D.T., M.P.; funding acquisition: A.F., K.A.G.; investigation: A.F., D.W.B., L.R.I., L.D.D., C.N.H., Q.Z., S.B.H., A.Y.M., A.G., K.N., K.A.G.; methodology: A.F., D.W.B., S.B.D., L.R.I., A.A.S., N.F.R., R.J.I.; project administration: A.F.,

K.A.G.; software: A.F., D.W.B., L.R.I., L.D.D., C.N.H., Q.Z., A.G.; supervision: A.F. K.A.G.; validation: A.F., D.W.B., L.D.D., A.G., Q.Z., S.M.D., L.R.I.; visualization: D.W.B., S.B.H., P.V., S.A.B., A.Y.M., C.N.H., R.J.I., A.A.S., N.F.R.; writing – original draft: A.F., D.W.B., C.N.H., S.U.B., K.A.G.; writing – review & editing: A.F., D.W.B., C.N.H., K.A.G., T.M.M., M.P.

Declaration of Interest

The authors declare no competing interests.

References

1. Aupperle, R. L. & Paulus, M. P. Neural systems underlying approach and avoidance in anxiety disorders. *Dialogues in clinical neuroscience* **12**, 517–31 (2010).
2. Crittenden, J. R. & Graybiel, A. M. Basal Ganglia disorders associated with imbalances in the striatal striosome and matrix compartments. *Front Neuroanat* **5**, 59 (2011).
3. Friedman, A. *et al.* Chronic Stress Alters Striosome-Circuit Dynamics, Leading to Aberrant Decision-Making. *Cell* **171**, 1191-1205.e28 (2017).
4. Graybiel, A. M. & Matsushima, A. Striosomes and Matrisomes: Scaffolds for Dynamic Coupling of Volition and Action. *Annu. Rev. Neurosci.* **46**, 359–380 (2023).
5. Nelson, A. B. & Kreitzer, A. C. Reassessing models of basal ganglia function and dysfunction. *Annu Rev Neurosci* **37**, 117–135 (2014).
6. Cox, J. & Witten, I. B. Striatal circuits for reward learning and decision-making. *Nat Rev Neurosci* **20**, 482–494 (2019).
7. Graybiel, A. M. & Ragsdale, C. W. Histochemically distinct compartments in the striatum of human, monkeys, and cat demonstrated by acetylthiocholinesterase staining. *Proc Natl Acad Sci U S A* **75**, 5723–5726 (1978).
8. Brimblecombe, K. R. & Cragg, S. J. The Striosome and Matrix Compartments of the Striatum: A Path through the Labyrinth from Neurochemistry toward Function. *ACS Chem Neurosci* **8**, 235–242 (2017).
9. Gerfen, C. R. Synaptic organization of the striatum. *J. Elec. Microsc. Tech.* **10**, 265–281 (1988).
10. Friedman, A. *et al.* Striosomes Mediate Value-Based Learning Vulnerable in Age and Huntington's Model. *Cell* **183**, 918-934.e49 (2020).
11. Friedman, A. *et al.* A Corticostriatal Path Targeting Striosomes Controls Decision-Making under Conflict. *Cell* **161**, 1320–1333 (2015).

12. Bloem, B. *et al.* Multiplexed action-outcome representation by striatal striosome-matrix compartments detected with a mouse cost-benefit foraging task. *Nat Commun* **13**, 1541 (2022).
13. Weglage, M. *et al.* Complete representation of action space and value in all dorsal striatal pathways. *Cell Rep* **36**, 109437 (2021).
14. Xiao, X. *et al.* A Genetically Defined Compartmentalized Striatal Direct Pathway for Negative Reinforcement. *Cell* **183**, 211-227.e20 (2020).
15. Yoshizawa, T., Ito, M. & Doya, K. Reward-predictive neural activities in striatal striosome compartments. *eNeuro* **5:e0367**, (2018).
16. Lazaridis, I. *et al.* Striosomes control dopamine via dual pathways paralleling canonical basal ganglia circuits. *Current Biology* S0960982224013381 (2024) doi:10.1016/j.cub.2024.09.070.
17. Jenrette, T. A., Logue, J. B. & Horner, K. A. Lesions of the Patch Compartment of Dorsolateral Striatum Disrupt Stimulus-Response Learning. *Neuroscience* **415**, 161–172 (2019).
18. Nadel, J. A. *et al.* Optogenetic stimulation of striatal patches modifies habit formation and inhibits dopamine release. *Sci Rep* **11**, 19847 (2021).
19. Nadel, J. A., Pawelko, S. S., Copes-Finke, D., Neidhart, M. & Howard, C. D. Lesion of striatal patches disrupts habitual behaviors and increases behavioral variability. *PLoS One* **15**, e0224715 (2020).
20. Prager, E. M. *et al.* Dopamine Oppositely Modulates State Transitions in Striosome and Matrix Direct Pathway Striatal Spiny Neurons. *Neuron* **108**, 1091-1102.e5 (2020).
21. Evans, R. C. *et al.* Functional Dissection of Basal Ganglia Inhibitory Inputs onto Substantia Nigra Dopaminergic Neurons. *Cell Reports* **32**, 108156 (2020).
22. Hong, S. *et al.* Predominant Striatal Input to the Lateral Habenula in Macaques Comes from Striosomes. *Curr Biol* **29**, 51-61.e5 (2019).

23. Rajakumar, N., Elisevich, K. & Flumerfelt, B. A. Compartmental origin of the striato-entopeduncular projection in the rat. *J Comp Neurol* **331**, 286–296 (1993).
24. Fujiyama, F. *et al.* Exclusive and common targets of neostriatofugal projections of rat striosome neurons: a single neuron-tracing study using a viral vector. *Eur J Neurosci* **33**, 668–677 (2011).
25. Schroll, H. & Hamker, F. H. Computational models of basal-ganglia pathway functions: focus on functional neuroanatomy. *Front. Syst. Neurosci.* **7**, (2013).
26. Crittenden, J. R. *et al.* Striosome-dendron bouquets highlight a unique striatonigral circuit targeting dopamine-containing neurons. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 11318–11323 (2016).
27. Barbera, G. *et al.* Spatially Compact Neural Clusters in the Dorsal Striatum Encode Locomotion Relevant Information. *Neuron* **92**, 202–213 (2016).
28. Azcorra, M. *et al.* Unique functional responses differentially map onto genetic subtypes of dopamine neurons. *Nat Neurosci* **26**, 1762–1774 (2023).
29. Wallace, M. L. *et al.* Genetically Distinct Parallel Pathways in the Entopeduncular Nucleus for Limbic and Sensorimotor Output of the Basal Ganglia. *Neuron* **94**, 138-152.e5 (2017).
30. Christoph, G. R., Leonzio, R. J. & Wilcox, K. S. Stimulation of the lateral habenula inhibits dopamine-containing neurons in the substantia nigra and ventral tegmental area of the rat. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **6**, 613–9 (1986).
31. Hong, S., Jhou, T. C., Smith, M., Saleem, K. S. & Hikosaka, O. Negative Reward Signals from the Lateral Habenula to Dopamine Neurons Are Mediated by Rostromedial Tegmental Nucleus in Primates. *Journal of Neuroscience* **31**, 11457–11471 (2011).
32. Lahiri, A. K. & Bevan, M. D. Dopaminergic Transmission Rapidly and Persistently Enhances Excitability of D1 Receptor-Expressing Striatal Projection Neurons. *Neuron* **106**, 277-290.e6 (2020).

33. Grillner, S. & Robertson, B. The basal ganglia downstream control of brainstem motor centres—an evolutionarily conserved strategy. *Current Opinion in Neurobiology* **33**, 47–52 (2015).
34. Balleine, B. W., Delgado, M. R. & Hikosaka, O. The Role of the Dorsal Striatum in Reward and Decision-Making. 5.
35. Lee, H. J. *et al.* Activation of Direct and Indirect Pathway Medium Spiny Neurons Drives Distinct Brain-wide Responses. *Neuron* **91**, 412–424 (2016).
36. Sanger, T. D. Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Networks* **2**, 459–473 (1989).
37. Kamiński, J. *et al.* Novelty-Sensitive Dopaminergic Neurons in the Human Substantia Nigra Predict Success of Declarative Memory Formation. *Current Biology* **28**, 1333-1343.e4 (2018).
38. Kim, S. & Bera, A. K. Scalar Measures of Volatility and Dependence for the Multivariate Models with Applications to Asian Financial Markets. *JRFM* **16**, 212 (2023).
39. Classical Microeconomics. in *On Classical Economics* 48–78 (Yale University Press, 2017). doi:10.12987/9780300185669-005.
40. Winterer, G. & Weinberger, D. R. Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. *Trends in Neurosciences* **27**, 683–690 (2004).
41. Miyamoto, Y., Katayama, S., Shigematsu, N., Nishi, A. & Fukuda, T. Striosome-based map of the mouse striatum that is conformable to both cortical afferent topography and uneven distributions of dopamine D1 and D2 receptor-expressing cells. *Brain structure & function* **223**, 4275–4291 (2018).
42. Hikosaka, O. *et al.* Direct and indirect pathways for choosing objects and actions. *Eur J Neurosci* **49**, 637–645 (2019).

43. Maltese, M., March, J. R., Bashaw, A. G. & Tritsch, N. X. Dopamine differentially modulates the size of projection neuron ensembles in the intact and dopamine-depleted striatum. *Elife* **10**, e68041 (2021).
44. Calabresi, P., Picconi, B., Tozzi, A., Ghiglieri, V. & Di Filippo, M. Direct and indirect pathways of basal ganglia: a critical reappraisal. *Nat Neurosci* **17**, 1022–1030 (2014).
45. Salinas, A. G., Davis, M. I., Lovinger, D. M. & Mateo, Y. Dopamine dynamics and cocaine sensitivity differ between striosome and matrix compartments of the striatum. *Neuropharmacology* **108**, 275–283 (2016).
46. Luo, Z., Volkow, N. D., Heintz, N., Pan, Y. & Du, C. Acute Cocaine Induces Fast Activation of D1 Receptor and Progressive Deactivation of D2 Receptor Striatal Neurons: *In Vivo* Optical Microprobe $[Ca^{2+}]_i$ Imaging. *J. Neurosci.* **31**, 13180–13190 (2011).
47. Oyama, K., Hernádi, I., Iijima, T. & Tsutsui, K.-I. Reward Prediction Error Coding in Dorsal Striatal Neurons. *J. Neurosci.* **30**, 11447–11457 (2010).
48. Cavedini, P. *et al.* Decision-making heterogeneity in obsessive-compulsive disorder: ventromedial prefrontal cortex function predicts different treatment outcomes. *Neuropsychologia* **40**, 205–211 (2002).
49. Cohen, S. E., Zantvoord, J. B., Wezenberg, B. N., Bockting, C. L. H. & Van Wingen, G. A. Magnetic resonance imaging for individual prediction of treatment response in major depressive disorder: a systematic review and meta-analysis. *Transl Psychiatry* **11**, 168 (2021).
50. Saperia, S. *et al.* Modeling Effort-Based Decision Making: Individual Differences in Schizophrenia and Major Depressive Disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **8**, 1041–1049 (2023).
51. Wright, A. G. C. & Simms, L. J. Stability and fluctuation of personality disorder features in daily life. *Journal of Abnormal Psychology* **125**, 641–656 (2016).

52. de Siqueira, A. S. S. *et al.* Decision Making assessed by the Iowa Gambling Task and Major Depressive Disorder A systematic review. *Dement Neuropsychol* **12**, 250–255 (2018).
53. Cousijn, J. *et al.* Individual differences in decision making and reward processing predict changes in cannabis use: a prospective functional magnetic resonance imaging study. *Addict Biol* **18**, 1013–1023 (2013).
54. Horn, S. R. & Feder, A. Understanding Resilience and Preventing and Treating PTSD. *Harv Rev Psychiatry* **26**, 158–174 (2018).
55. Pickens, C. L., Golden, S. A., Adams-Deutsch, T., Nair, S. G. & Shaham, Y. Long-lasting incubation of conditioned fear in rats. *Biol. Psychiatry* **65**, 881–6 (2009).
56. Fredriksson, I. *et al.* Orbitofrontal cortex and dorsal striatum functional connectivity predicts incubation of opioid craving after voluntary abstinence. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2106624118 (2021).
57. Gurney, K., Prescott, T. J. & Redgrave, P. A computational model of action selection in the basal ganglia. II. Analysis and simulation of behaviour. *Biol Cybern* **84**, 411–423 (2001).
58. Möller, M., Manohar, S. & Bogacz, R. Uncertainty–guided learning with scaled prediction errors in the basal ganglia. *PLoS Comput Biol* **18**, e1009816 (2022).
59. Morita, K. & Kato, A. Striatal dopamine ramping may indicate flexible reinforcement learning with forgetting in the cortico-basal ganglia circuits. *Front. Neural Circuits* **8**, (2014).
60. Dabney, W. *et al.* A distributional code for value in dopamine-based reinforcement learning. *Nature* **577**, 671–675 (2020).
61. Bar-Gad, I., Morris, G. & Bergman, H. Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. *Progress in Neurobiology* **71**, 439–473 (2003).
62. Berns, G. S. & Sejnowski, T. J. A Computational Model of How the Basal Ganglia Produce Sequences. *Journal of Cognitive Neuroscience* **10**, 108–121 (1998).

63. Garr, E. Contributions of the basal ganglia to action sequence learning and performance. *Neuroscience & Biobehavioral Reviews* **107**, 279–295 (2019).
64. Reiner, A. Corticostriatal projection neurons – dichotomous types and dichotomous functions. *Front. Neuroanat.* **4**, (2010).
65. Lanciego, J. L., Luquin, N. & Obeso, J. A. Functional neuroanatomy of the basal ganglia. *Cold Spring Harb Perspect Med* **2**, a009621 (2012).
66. Martel, A.-C. & Galvan, A. Connectivity of the corticostriatal and thalamostriatal systems in normal and parkinsonian states: An update. *Neurobiology of Disease* **174**, 105878 (2022).
67. Eblen, F. & Graybiel, A. M. Highly restricted origin of prefrontal cortical inputs to striosomes in the macaque monkey. *J Neurosci* **15**, 5999–6013 (1995).
68. McGregor, M. M. *et al.* Functionally Distinct Connectivity of Developmentally Targeted Striosome Neurons. *Cell Rep* **29**, 1419-1428.e5 (2019).
69. Gerfen, C. R. The neostriatal mosaic: compartmentalization of corticostriatal input and striatonigral output systems. *Nature* **311**, 461–4 (1984).
70. Smith, J. B. *et al.* Genetic-Based Dissection Unveils the Inputs and Outputs of Striatal Patch and Matrix Compartments. *Neuron* **91**, 1069–1084 (2016).
71. Hong, S. & Hikosaka, O. Diverse sources of reward value signals in the basal ganglia nuclei transmitted to the lateral habenula in the monkey. *Front Hum Neurosci* **7**, 778 (2013).
72. Hong, S. & Hikosaka, O. The Globus Pallidus Sends Reward-Related Signals to the Lateral Habenula. *Neuron* **60**, 720–729 (2008).
73. Weglage, M. *et al. Sst+ GPi Output Neurons Provide Direct Feedback to Key Nodes of the Basal Ganglia and Drive Behavioral Flexibility*. http://biorxiv.org/lookup/doi/10.1101/2022.03.16.484460 (2022) doi:10.1101/2022.03.16.484460.

74. Gonçalves, L., Sego, C. & Metzger, M. Differential projections from the lateral habenula to the rostromedial tegmental nucleus and ventral tegmental area in the rat. *J of Comparative Neurology* **520**, 1278–1300 (2012).
75. Jhou, T. C., Fields, H. L., Baxter, M. G., Saper, C. B. & Holland, P. C. The Rostromedial Tegmental Nucleus (RMTg), a GABAergic Afferent to Midbrain Dopamine Neurons, Encodes Aversive Stimuli and Inhibits Motor Responses. *Neuron* **61**, 786–800 (2009).
76. Smith, R. J., Vento, P. J., Chao, Y. S., Good, C. H. & Jhou, T. C. Gene expression and neurochemical characterization of the rostromedial tegmental nucleus (RMTg) in rats and mice. *Brain Struct Funct* **224**, 219–238 (2019).
77. Jaquins-Gerstl, A., Nesbitt, K. M. & Michael, A. C. In vivo evidence for the unique kinetics of evoked dopamine release in the patch and matrix compartments of the striatum. *Anal Bioanal Chem* (2021) doi:10.1007/s00216-021-03300-z.
78. Münte, T. F. *et al.* The human globus pallidus internus is sensitive to rewards – Evidence from intracerebral recordings. *Brain Stimulation* **10**, 657–663 (2017).
79. Stephenson-Jones, M. *et al.* A basal ganglia circuit for evaluating action outcomes. *Nature* **539**, 289–293 (2016).
80. Matsumoto, M. & Hikosaka, O. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* **447**, 1111-U11 (2007).
81. Lee, H. & Hikosaka, O. Lateral habenula neurons signal step-by-step changes of reward prediction. *iScience* **25**, 105440 (2022).
82. Stopper, C. M. & Floresco, S. B. What's better for me? Fundamental role for lateral habenula in promoting subjective decision biases. *Nat Neurosci* **17**, 33–35 (2014).
83. Vento, P. J., Burnham, N. W., Rowley, C. S. & Jhou, T. C. Learning From One's Mistakes: A Dual Role for the Rostromedial Tegmental Nucleus in the Encoding and Expression of Punished Reward Seeking. *Biological Psychiatry* **81**, 1041–1049 (2017).

84. Fiorillo, C. D., Tobler, P. N. & Schultz, W. Discrete Coding of Reward Probability and Uncertainty by Dopamine Neurons. *Science* **299**, 1898–1902 (2003).
85. Matsumoto, M. & Hikosaka, O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* **459**, 837–841 (2009).
86. Gan, J. O., Walton, M. E. & Phillips, P. E. M. Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nat Neurosci* **13**, 25–27 (2010).
87. Bromberg-Martin, E. S., Matsumoto, M. & Hikosaka, O. Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. *Neuron* **68**, 815–834 (2010).
88. Kim, H. R. *et al.* A Unified Framework for Dopamine Signals across Timescales. *Cell* **183**, 1600-1616.e25 (2020).
89. Long, C. *et al.* Constraints on the subsecond modulation of striatal dynamics by physiological dopamine signaling. *Nat Neurosci* **27**, 1977–1986 (2024).
90. Samejima, K., Ueda, Y., Doya, K. & Kimura, M. Representation of action-specific reward values in the striatum. *Science* **310**, 1337–40 (2005).
91. Seo, M., Lee, E. & Averbeck, B. B. Action Selection and Action Value in Frontal-Striatal Circuits. *Neuron* **74**, 947–960 (2012).
92. Parker, J. G. *et al.* Diametric neural ensemble dynamics in parkinsonian and dyskinetic states. *Nature* **557**, 177–182 (2018).
93. Peak, J., Chieng, B., Hart, G. & Balleine, B. W. Striatal direct and indirect pathway neurons differentially control the encoding and updating of goal-directed learning. *eLife* **9**, e58544 (2020).
94. Sun, F. *et al.* Next-generation GRAB sensors for monitoring dopaminergic activity in vivo. *Nat Methods* **17**, 1156–1166 (2020).

Table 1: Terminology

Term	Definition
GPi	Globus pallidus internus
GPe	Globus pallidus externus
LHb	Lateral habenula
RMTg	Rostromedial tegmental nucleus
daSNC	dopaminergic neurons of the substantia nigra compacta
FSI	Striatal fast-spiking Interneuron
sSPN	Striosomal striatal projection neuron
mSPN	Matrix striatal projection neuron
dsSPN	Direct pathway sSPN
isSPN	Indirect pathway sSPN
dmSPN	Direct pathway mSPN
imSPN	Indirect pathway mSPN
BG	Basal ganglia
Decision-dimension	An axis of the coordinate system with which SPNs (dsSPNs, isPSNs, dmSPNs, imSPNs) process cortical activity. Subpopulations of SPNs (for each dsSPNs, isPSNs, dmSPNs, imSPNs) encode data along different decision-dimensions. Cortical activity is linearly mapped to this basis of decision-dimensions such that the activity of a single cortical neuron no longer is encoded by a single SPN. Certain decision-dimensions might correspond more predominantly, for instance, to reward level, cost level, or novelty level, as encoded across multiple cortical neurons. Decision-dimensions are modeled as the principal components of cortical activity.
Decision-space	The subspace produced by the decision-dimensions which are selected by the circuit to be used during a decision. Decision-space is formed when dopamine releases to mSPNs (dmSPNs and imSPNs), signaling that certain decision-dimensions are important and others are unimportant, and therefore can be excluded from the subspace.
Action value	Value assigned by the circuit to an action
Inaction value	Value assigned by the circuit for refraining from an action
Prediction error	The difference between expected and observed information along a data axis (for instance, reward prediction error, punishment prediction error, or novelty prediction error).
Circuit activity	The set of average activities of each circuit element during a decision
Advantage	The degree to which a circuit activity is preferred. Used in our analysis of the change in circuit activity between trials.

Table 2: Evidence for the connectivity used in our model.

Connectivity	Evidence
cortex→sSPN/mSPN	• Research has found that nearly, if not all cortical regions project to the striatum[64,65], however only a subset of cortical areas have been determined to be compartment specific[4,41], with examples of those listed below. • Evidence that both the striosomes and matrix receive input from sensorimotor, limbic, and associative regions[66]. • Evidence that several regions (prelimbic cortex, infralimbic cortex, posterior orbitofrontal cortex, insular cortex) project more to striosomes than matrix[2,67]. • Evidence that several regions (primary motor cortex) project more to matrix than to striosomes[68].
cortex→FSI→sSPN/mSPN	• Evidence for FSI connections to both striosomes and matrix[3,10,11].
sSPN (more so than mSPN) →daSNC	• Range of evidence from primates, rats[24,69], and mice[68] suggesting a stronger connection striosome connection than matrix. • One study suggests that matrix also projects to daSNC[70].
sSPN (more so than mSPN) →GPi→LHb→RMTg→daSNc	• Range of evidence supporting the striosome to EP (non-primate GPi correlate) connection in rats[23,24]. • Evidence supporting the striosome to GPi connection in primates[71]. • Evidence of the EP to LHb connection in rats[72]. • Evidence that striosomes more so than matrix drive striatal influence on LHb activity[22], potentially through pallidal regions[73]. • Evidence in primates[72], rats[74] , and mice[73] of the LHb to RMTg (also called tVTA) connection. • Evidence that the RMTg projects to daSNc[75,76].
sSPN (more so than mSPN) →GPe→daSNC	• Evidence of the striosome to GP (non-primate GPe correlate) but not matrix to GP connection in mice[16] and rats[24]. • Evidence of a striosome (more so than matrix) to GPe to daSNC pathway[16].
daSNC→sSPN/mSPN	• Evidence that dopamine is released to both compartments, with faster[77] and more[8,45] dopamine release to striosomes than matrix.

Table 3: Criteria used in Tables 4-8 to test our decision-space model and compare it to alternative models of the circuit.

Modeled Functional role of a circuit region.	**Expectation per the model.**
sSPNs influence the priority of decision-dimensions, thereby affecting decision-space. **(decision-space model)**	**1.A.** During a high-dimensional decision-space, choice more closely aligns to experimental inputs (e.g. chocolate milk level, light brightness) during difficult tasks (e.g. consideration of both rewards and costs).
	1.B. A low-dimensional decision-space is often formed from decision-dimensions which are commonly used during a decision-making task, for instance information about reward in rodents that are trained to respond to reward cues.
	1.C. Cortical data is mapped orthogonally and continuously onto SPNs.
sSPNs encode conflict. **(our conflict model)**	**2.A.** sSPN activity scales with conflict between two features like reward and cost.
	2.B. Changes to conflict, revealed by sSPN activity, alter choice.
	2.C. Changes to sSPN signals are greatest when conflict is introduced.
sSPNs encode subjective values. **(our subjective value model)**	**3.A.** sSPN activity scales (possibly directly or inversely) with subjective value of stimuli, likely roughly tracking reward minus cost.
	3.B. Higher subjective value, reflected by sSPN activity, leads to increased selection of an offer.
	3.C. Changes to sSPN activity are strongest at the time during scenarios where cues are associated with subjective values.
sSPNs encode prediction errors. **(our prediction error model)**	**4.A.** sSPN activity should change proportionally and continuously to the difference between expected and actual reward and cost at each time step.
	4.B. As a task is learned, the trend in sSPN activity over time changes as earlier cues become associated with later outcomes.
sSPNs encode actions. **(our actions model)**	**5.A.** Different neurons would encode different actions
	5.B. Activity of sSPN subpopulations should scale directly or inversely with the predisposition of action execution.
	5.C. Changes to the signals of action-encoding subpopulation would be greatest prior to or during action execution.
GPi during decision-space formation. **(decision-space model)**	**6.A.** Changes to activity along a decision-dimension is reflected in GPi activity.
	6.B. A given GPi neuron encodes data across multiple decision-dimensions.
	6.C. Activation of GPi causes more decision-dimensions to be incorporated into the decision-space while inactivation causes fewer decision-dimensions to be used.
LHb and RMTg optimizing or modifying decision-space. **(decision-space model)**	**7.A.** Active LHb (or RMTg) leads to choice reflective of reduction in dimensionality of the decision-space and vice versa.
	7.B. LHb (or RMTg) is active during times when lower-dimensional decision-spaces are beneficial to decision-making.
	8.A. There exist subpopulations of daSNC neurons that encode information along an orthogonal axis of information.

Dopaminergic neurons of the SNc during decision-space formation. **(decision-space model)**	**8.B.** Activity in one daSNC subpopulation only affects the subpopulation of SPNs corresponding to one decision-dimension.
Direct and indirect pathways alter decision-space formation. **(decision-space model)**	**9.A.** Higher dopamine leads to lower dimensionality of the direct pathway decision-space while lower dopamine leads to higher dimensionality, and vice versa for the indirect pathway.
	9.B. Direct pathway mSPNs promote actions while indirect pathway mSPNs aid in action suppression.
	9.C. Subpopulations of mSPNs encode data along decision-dimensions orthogonally and continuously over time.

729
730

Table 4: Testing alignment of the decision-space model and other models to a selection of the experimental sSPN literature.

	Criterion	**Friedman et al. (2015)[11] (I)**	**Friedman et al. (2020)[10] (II)**	**Bloem et al. (2022)[12] (III)**	**Xiao et al. (2020)[14] (IV)**	**Weglage et al. (2021)[13] (V)**
Decision-space model	1.A	✓	✓	∅	≈	∅
	1.B	∅	✓	∅	∅	∅
	1.C	∅	∅	✓	✓	✓
Conflict model	2.A	✓	×	×	×	∅
	2.B	✓	∅	∅	∅	∅
	2.C	✓	×	×	×	∅
Subjective value model	3.A	×	✓	×	×	×
	3.B	✓	✓	∅	✓	×
	3.C	✓	✓	✓	✓	×
Prediction error model	4.A	×	×	✓	×	×
	4.B	∅	×	✓	✓	∅
Actions model	5.A	∅	∅	✓	✓	≈
	5.B	✓	✓	✓	✓	≈
	5.C	≈	✓	✓	✓	≈

✓ -- aligned with criterion
≈ -- somewhat aligned to criterion
× -- not aligned with criterion
∅ -- experiment does not test criterion

1.A.I. Striosome activation led to less consistent choice in tasks that required processing of both reward and cost, while sSPN inhibition led to more consistent choice in tasks that required processing of both reward and cost. Decision-making was less affected by sSPN manipulation in tasks that required processing either reward or cost, but not both. Meanwhile, in the absence of manipulation, the rodents had less active sSPNs during tasks that required processing both reward and cost. These results may suggest that a higher-dimensional decision-space (formed at low striosome activity) is associated with processing of multiple informational dimensions in a consistent way.

1.A.II. Rodents that best learned the reward/cost cue discrimination task had high sSPN activity after the reward cue and low sSPN activity after the cost cue. The rodents that learned less well had similar activity between the tasks. This may suggest that the most consistent choices, made by the rodents that learned, were formed using a context-dependent decision-space.

1.A.IV. Inhibition of Tfz1 neurons, which were shown to encode either reward or cost independently, decreased accuracy and correct rejection rate in cost trials. This may suggest a

change to the priority assigned to the reward and cost decision-dimensions. Notably, the direction of the effect is opposite what the model expects, if only cost-related SPNs were inhibited. However, Tfz1 neurons that responded to rewards and costs were inhibited simultaneously, making for a less clear model prediction on the direction of the effect.

1.B.II. More rodents successfully learned the reward task than the cost task, and this was reflected in fewer rodents forming a reduced sSPN activity and reaching a putative high-dimensional decision-space. This may suggest that most rodents formed a one-dimensional reward-related decision-space and only the rodents that learned formed, in the appropriate context, a two-dimensional reward-and-cost-related decision-space.

1.C.III. In the probabilistic bandit task, dynamic changes along the orthogonal information axes of reward and cost led to proportional changes in sSPNs that resembled prediction errors. Some of the identified neurons responded only to reward or only to cost. This may suggest that those neurons are members of subpopulations corresponding to a reward-predominant decision-dimension or a cost-predominant DM-dimension.

1.C.IV. Tfz1-expressing sSPNs showed evidence of encoding reward or cost dimensions independently (although it was not confirmed that Tfz1 neurons were responding to cortical signals). During a real-time place preference task, different neurons were activated by either reward or cost individually both during administration of the reward or cost and at a cue associated with the reward or cost. This may suggest encoding of information along decision-dimensions, including related information observed from separate stimuli.

1.C.V. During a multiphase task that required attention to many stimuli and strategies, sSPN and mSPN activities in general followed somewhat similar activity patterns over time. The most significant indicator of SPN activity was task phase. This may suggest a continuous mapping, with similar weights between sSPN and mSPN, of cortical information relevant to the current phase of the task onto the decision-space in the striatum.

2.A.I. Mean sSPN activity was significantly different in the task where there was the most conflict than in the other tasks.

2.A.II-IV. Conflict was not introduced experimentally, yet sSPN activity was different between tasks.

2.B.I. Optogenetic manipulation of striosomes, which may alter the level of encoded cost/reward conflict, altered choice.

2.C.I. In the task, conflict is important to the rodent when it determines which option to select, which is when striosome activity rises. This may suggest that striosomes encode conflict.

2.C.II-IV. Conflict was not introduced experimentally, yet sSPN activity spiked when non-conflict stimuli were introduced.

3.A.I. sSPN activity did not scale, directly (more sSPN activity, more subjective value) nor inversely (more sSPN activity, less subjective value), with either the overall values of the options or with the difference between the value of the options on the right arm of the T-maze versus the left arm. sSPNs had lowest mean activity in the cost-benefit conflict task, which had a moderate difference between reward and cost. Both the benefit-benefit task, which had high reward and no

cost, and the cost-cost task, which had low reward and high cost, were performed with high sSPN activities.

3.A.II. sSPN activity was higher during the reward task than the cost task among rodents that learned the task. This may suggest that sSPN activity tracks subjective value of the presented stimuli, with the reward-related stimuli being assigned higher subjective value than the cost-related stimuli.

3.A.III. The activities of a sizable subpopulation of SPNs were found to track reward level and cost level separately, not together as our subjective value model would expect. Meanwhile, neurons that tracked both reward and cost had activities that did not scale with reward minus cost, suggesting encoding of information other than the subjective value of the stimuli.

3.A.IV. Separate SPN subpopulations were found to track reward or cost level, suggesting encoding of reward and cost separately, rather than together as our subjective value model would expect.

3.A.V. There was not a strong relationship between rewards minus costs presented in the different task phases and sSPN activity.

3.B.I. The rodents approached the higher-reward option most in the task (cost-benefit conflict task) which had the lowest sSPN activity, suggesting a possible relationship between subjective value assigned to the left and right arms of the T-maze and sSPN activity.

3.B.II. The choice to lick was made more frequently when sSPN activity during the licking period was higher.

3.B.IV. Stimulation and inhibition of Tfz1-expressing neurons led to opposite effects on decision-making, suggesting that manipulation may change subjective values assigned to stimuli.

3.B.V. There was not a strong relationship between sSPN activity and choice in the multiphase task.

3.C.I. Across all tasks, sSPN activity ramped during the period of the decision where a choice was made as to which offer to approach, which is also likely the period of the task that most requires an assignment of subjective value to environmental stimuli. This suggests that perhaps sSPNs play a role in assigning subjective value.

3.C.II. sSPN activity rose during the periods when the rodents chose whether to lick, which is the period of the task that likely most requires an assignment of subjective value to environmental stimuli. This suggests that perhaps sSPNs play a role in assigning subjective value.

3.C.III. sSPN activity was the highest at the cue and during the outcome period, two intervals when it is likely important to assign subjective value to stimuli. This suggests that perhaps sSPNs play a role in assigning subjective value.

3.C.IV. sSPN activity rose during the cue and administration periods of the Pavlovian conditioning task, both of which are likely important for assignment of subjective value. This suggests that perhaps sSPNs play a role in assigning subjective value.

3.C.V. sSPNs did not have exceptionally high nor low activities during periods of the multiphase task in which we might expect subjective value assignment to be important. sSPN activity was higher during a locomotion phase, for example, than during a phase where a reward was presented.

4.A.I. The rodents in the experiments were overtrained, task types were randomized, and each task type had different levels of reward and/or cost. Therefore, they might be expected to experience prediction error when they entered the maze and become aware of the task type. Prediction upon entering the maze should be roughly equivalent to the average value (reward minus cost) of all tasks. By this logic, there is positive prediction error in tasks where the expected outcome was greater than the expected outcome upon entering the maze, and vice versa. Thus, the largest positive prediction error likely occurred during the benefit-benefit task, followed by the cost-benefit non-conflict task, followed by the cost-benefit conflict task, followed by the cost-cost task. Experimental sSPN activities were not ordered like this.

4.A.II. Theoretically, we would expect, post-learning, a prediction error at the tone because the value of the session changes at this point. Yet sSPN activity had little change at the tone post-learning.

4.A.III. The authors demonstrated that separate populations track prediction error directly and inversely.

4.A.IV. Theoretically, prediction error moves from the administration period to the cue during a Pavlovian task. This is not what was observed: sSPN activity spiked during the administration period just as much after learning as before.

4.A.V. sSPN activity had little change at periods of the task when prediction errors were introduced experimentally.

4.B.II. sSPN activity over time did change with learning but spikes in activity did not develop at times (e.g. the cue) we would expect post-learning prediction errors to develop.

4.B.III. sSPN activity tracked prediction error more closely at the cue after learning than before. This suggests that sSPNs may be encoding prediction error and learning the association between the cue and the outcome.

4.B.IV. The activities of individual SPN increased in magnitude at the cue after learning in the Pavlovian conditioning task (although there was not, as our prediction error model would expect, a corresponding reduction during administration). In the active avoidance task, failure-responding sSPNs increased activity upon punishment delivery.

5.A.III. Different neurons were found to respond to reward versus cost. This may suggest that different neurons were involved in encoding actions planned in response to the different stimuli.

5.A.IV. Different neurons were active in response to reward versus cost. This may suggest that different neurons were involved in encoding of actions planned in response to the different stimuli.

5.A.V. Neurons encoded actions in task-specific contexts but remapped between tasks.

5.B.I. The optogenetic manipulation caused altered actions (turn right versus turn left), suggesting that perhaps stimulation or inhibition leads to upweighting of one action versus another.

5.B.II. Different actions (licking versus non-licking) corresponded to different mean sSPN activity. This may suggest that sSPNs encode the action of licking.

5.B.III. sSPN subpopulations changed in response to unexpected stimuli, perhaps to recalculate actions to take.

5.B.IV. The reward-active sSPNs were most active during the outcome period and less so during the cue. This could be a sign of a potential action linked to the stimuli.

5.B.V. sSPN subpopulations were identified but their activities were only related to actions on a task-by-task basis.

5.C.I. For many trials, sSPN activity increased during the period when the animal made a choice between turning left versus turning right. However, sSPN activity was roughly constant throughout the choice periods of cost-benefit conflict trials, and the actions model would expect a ramping of activity here, too.

5.C.II. sSPN activity increased during the period when the animal made a choice whether or not to lick. This may suggest that sSPNs encode the action of licking.

5.C.III. sSPN subpopulations responded to the prediction error. Through the lens of the action model, perhaps sSPNs are revising the potential actions that will be initiated.

5.C.IV. sSPN activity was positively correlated with running velocity, and sSPN activity ramped along with licking bouts throughout the cue and outcome periods. This suggests that perhaps sSPNs encode the action of running.

5.C.V. sSPN subpopulations were most active when certain actions (e.g. turn direction) occurred, but between tasks these were not responsive during the same actions.

Table 5: Testing the alignment of the decision-space model to a selection of the experimental literature on GPi.

Criterion	**Weglage et al. (2022)[73] (I)**	**Munte et al. (2017)[78] (II)**	**Stephenson-Jones et. al (2016)[79] (III)**
6.A	∅	✓	∅
6.B	∅	∅	✓
6.C	✓	∅	∅

✓ -- aligned with criterion
≈ -- somewhat aligned to criterion
× -- not aligned with criterion
∅ -- experiment does not test criterion

6.A.II. Level of reward correlated with GPi activity. This may suggest that GPi activity scales up or down depending on the level of information along a reward-related decision-dimension.

6.B.III. Individual LHb-projecting GPi neurons were both excited by punishment-predicting cues and the punishment itself and were inhibited by rewards and their associated cues. This may suggest that information along two decision-dimensions, one reward-related and one-cost related, is encoded by the same GPi neurons. Further, the opposite response of GPi neurons to reward and cost lends support for our choice to differentially weight GPi inputs from sSPN subpopulations corresponding to different decision-dimensions.

6.C.I. An identified subpopulation of LHb-projecting GPi affected the profile of choices made, aligning with the functional role of the GPi in our model. Decreased activity of these neurons was associated with increased commitment to actions, which may correspond to effective formation of a decision-space. This would align with the hypothesis of the model that lower GPi activity leads to a higher-dimensional decision-space.

Table 6: Testing the alignment of the decision-space model to a selection of the experimental literature on LHb and RMTg.

Criterion	**Matsumoto & Hikosaka (2007)[80] (I)**	**Lee & Hikosaka (2022)[81] (II)**	**Stopper & Floresco (2014)[82] (III)**	**Vento et al. (2017)[83] (IV)**
7.A	∅	∅	≈	✓
7.B	✓	✓	∅	∅

✓ -- aligned with criterion
≈ -- somewhat aligned to criterion
× -- not aligned with criterion
∅ -- experiment does not test criterion

7.A.III. LHb inactivation led subjects to change their choice during a probabilistic discounting task to accept a large, risky reward over a smaller, safe reward. As our model expects, LHb inactivation played an important role in affecting decisions that required multiple decision-dimensions. It is expected, however, that reduced LHb activity produces enhanced adherence to any decision-dimensions required to perform the task. The subjects with inactive LHb would appear to be incorporating fewer, not more, decision-dimensions into their choices. One possible explanation is that LHb inactivation led to an overwhelming increase in dimensionality of the decision-space that reduced focus on a few important decision-dimensions, such as reward.

7.A.IV. RMTg selectively altered decisions, primarily in response to cost. This may suggest that RMTg inactivation led to choices with less adherence to an important cost-related decision-dimension.

7.B.I. LHb activation led to suppression of dopaminergic signaling among daSNC neurons. LHb was active at times when it may not have been beneficial to construct a decision-space involving a reward-related decision-dimension (when no reward was presented) but inactive when it may have been beneficial to construct a decision-space using a reward-related decision-dimension (when reward was presented).

7.B.II. LHb was found to alter its activity depending on situational context. LHb was most active at times when it may not have been beneficial to construct a decision-space involving a reward-predominant decision-dimension (when it was indicated that minimal reward would be available or when less than expected reward was presented) and at times when evaluation of data along decision-dimensions may have been less necessary (the uncontrollable tasks).

Table 7: Testing the alignment of the decision-space model to a selection of the experimental literature on daSNC.

Criterion	Fiorillo et al. (2003)[84] (I)	Matsumoto & Hikosaka (2009)[85] (II)	Gan et al. (2010)[86] (III)	Bromberg-Martin et al. (2010)[87] (IV)	Kim et al. (2020)[88] (V)	Long et al. (2024)[89] (VI)
8.A	✓	✓	✓	✓	✓	∅
8.B	∅	∅	∅	∅	∅	≈

✓ -- aligned with criterion
≈ -- somewhat aligned to criterion
× -- not aligned with criterion
∅ -- experiment does not test criterion

8.A.I. daSNC neurons responded differently during the cue and during the outcome period depending on the likelihood of a cue predicting a reward outcome. This may suggest that a subpopulation of daSNC neurons encodes information along a reward-related decision-dimension, and probabilistic inputs are reflected continuously over time as the information along the decision-dimension is updated.

8.A.II. Two daSNC populations responded very differently to rewarding or aversive stimuli and a third group was non-responsive. This may support the tenet of the decision-space model that different daSNC subpopulations correspond to different decision-dimensions, some of which might be related to reward information, some to cost information, and some to neither reward nor cost information.

8.A.III. The activities of recorded dopamine neurons showed more resemblance to reward levels than to overall utility. This may support the tenet of the decision-space model that different dopamine subpopulations correspond to different decision-dimensions, some of which are related to reward information, and that dopamine neurons encode data along decision-dimensions, not an overall value function.

8.A.IV. Subpopulations of daSNC neurons that encoded value were excited by rewarding information while salience neurons were excited by both rewarding and aversive cues. This may support the tenet of the decision-space model that different daSNC subpopulations correspond to different decision-dimensions, some of which are related to reward information and some to other information.

8.A.V. Dopamine changed in response to altered proximity to reward. This may support the architecture of the decision-space model, where changes to reward information are captured in an sSPN subpopulation related to reward, then passed to a corresponding daSNC subpopulation.

8.B.VI. VTA cells were optogenetically inhibited or excited as ventral striatal neurons were recorded. Several findings are particularly relevant to the decision-space model: 1) A subpopulation of the striatal neurons responded to reward, and the activities of these neurons

correlated with the VTA neurons. 2) 8% of all SPNs (4% above control, both non-reward-responding and reward-responding) had altered activities when VTA was inhibited. 3) The physical location of the SPNs that had altered activities had significantly distinct locations. Finding 1 may support the decision-space model, where mSPNs and sSPNs receive somewhat similar cortical inputs and sSPNs influence daSNC activity. A proportion of the reward-responding SPNs may encode a reward-related decision-dimension (the others may encode other information about the task and the reward administration). Further, per Finding 2, only a proportion of SPNs were affected by the reward-induced daSNC activity, supporting the selective release of dopamine to a reward-related SPN subpopulation in the model. Per Finding 3, the SPNs that did change their activities had spatial organization, supporting the assumption of our model that SPN subpopulations corresponding to decision-dimensions are organized spatially. The experiments, however, were primarily conducted on VTA neurons, not daSNC neurons, and a decoder did not accurately discriminate inhibition from control trials based on SPN spiking. This might be because a reward-related decision-space would require the modulation of more than 8% of neurons. Per our model, a decision-space would be formed in cases where sufficient dopamine was released to modulate a larger percentage (but not all SPNs). Indeed, when VTA neurons were manipulated to release more dopamine than they ordinarily did during the head-fixed licking task, up to 37% SPNs responded and a decoder successfully linked firing rates to the task, suggesting the formation of a decision-space. It may be that a decision-space only forms in certain tasks, for instance perhaps in tasks that require action selection, a hypothesis that might be supported by the findings of Samejima et al.[90] and Seo et al.[91] Our model provides a framework for task-dependent dopamine release to be studied in future work.

Table 8: Interpreting a selection of the experimental literature on the direct versus indirect pathways through the decision-space model.

Criterion	**Parker et al. (2018)[92] (I)**	**Peak et al. (2020)[93] (II)**	**Maltese et al. (2021)[43] (III)**	**Barbera et al. (2016)[27] (IV)**
9.A	✓	∅	✓	✓
9.B	∅	✓	∅	✓
9.C	∅	∅	∅	✓

✓ -- aligned with criterion
≈ -- somewhat aligned to criterion
× -- not aligned with criterion
∅ -- experiment does not test criterion

9.A.I. The direct and indirect pathway were found to be typically coactivated, and dopamine depletion differentially altered direct versus indirect pathway dynamics. This may support the tenet of our model that a decision-space is formed in each pathway during a decision and the two decision-spaces are affected differently by dopamine.

9.A.III. Increased dopamine release led to increased activation of dSPNs and decreased activation of iSPNs. Decreased dopamine release led to the opposite change. This supports the decision-space model, in which dopamine excites dmSPNs and inhibits imSPNs.

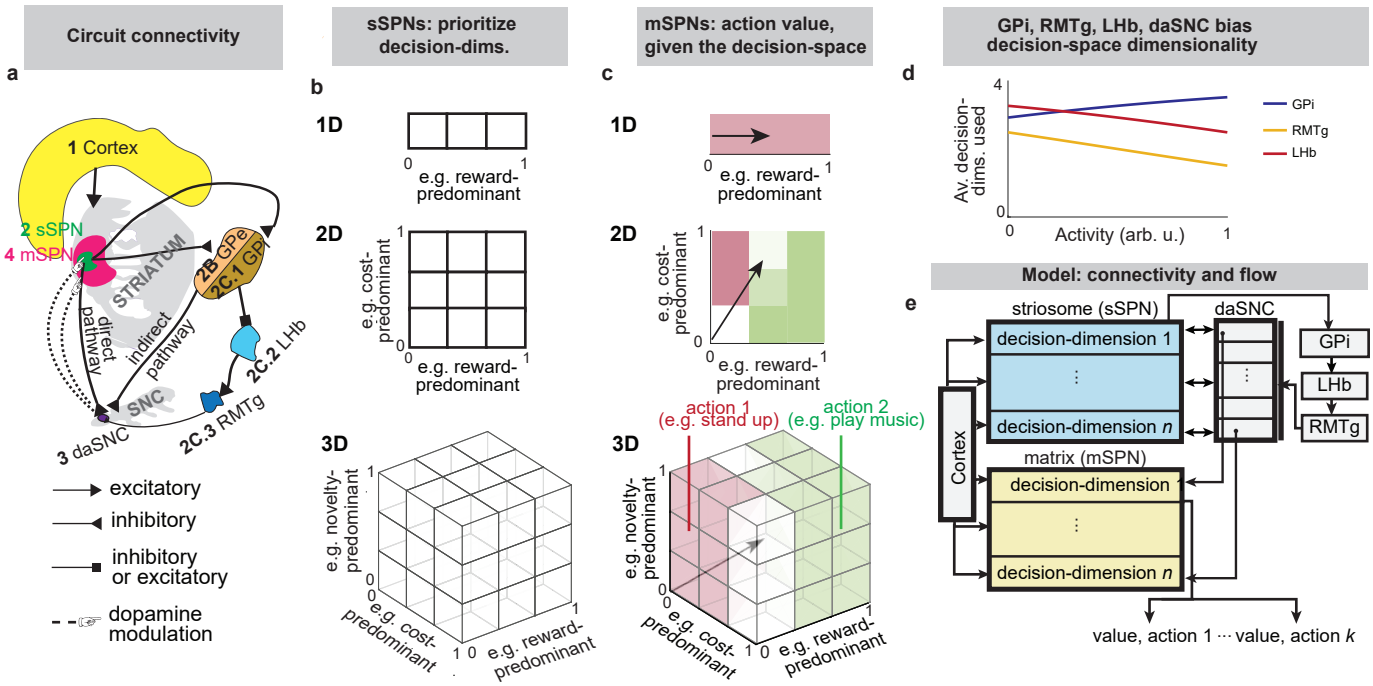
9.A.IV. Cocaine led to the increase in activity of a direct pathway subpopulation and simultaneously a decrease in activity of a neighboring indirect pathway subpopulation. This may suggest the simultaneous activation of a direct pathway decision-dimension and inactivation of a direct pathway decision-dimension.

9.B.II. Inhibition of dSPNs during learning led to blunted action associations, while inhibition of iSPNs led to a reduced ability to switch actions based on context. This supports the decision-space model, in which the direct pathway is involved with performing actions and the indirect pathway is involved in refraining from actions.

9.B.IV. Cocaine administration led to increased direct pathway activity, decreased indirect pathway activity, and more movement. This may suggest that the direct pathway more than the indirect pathway is associated with initiating actions.

9.C.IV. Subpopulations encoding data along decision-dimensions showed high intra-cluster synchrony that was stable across days, and inter-cluster synchrony was significantly lower. This may support the tenet of the decision-space model that similar information is encoded within proximate SPN subpopulations.

Figure 1



Model: functional roles

f

Cortex		input activity	eqs. 1,3
which dimensions?	sSPN	map to basis of decision-dimensions	eq. 1
	how many dimensions? GPi	decision-dimension combination	eq. 5
	how many dimensions? LHb	bias towards fewer decision-dimensions	eq. 5
	how many dimensions? RMTg	bias towards fewer decision-dimensions	eq. 5
	daSNC	final decision-dimension prioritization	eq. 2
mSPN		map to basis of decision-dimensions, action valuation	eq. 3

Model: example in real-life scenario

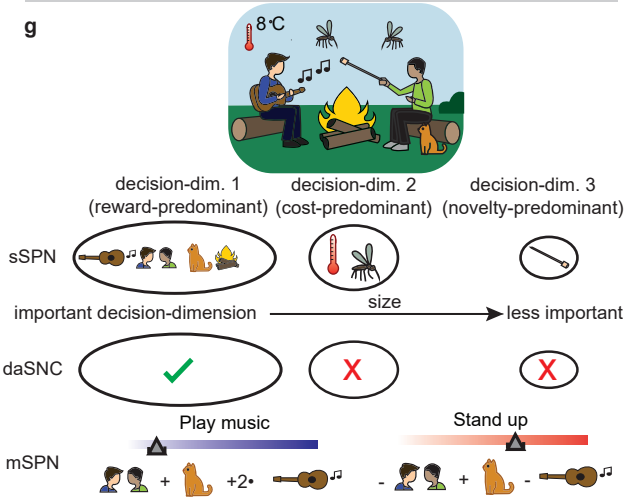


Figure 1 legend starts on next page

Fig. 1: The circuit defines a decision-space for action valuation.

a, The circuit of the cortex, striosomes (sSPN), globus pallidus externus (GPe), globus pallidus internus (GPi), lateral habenula (LHb), rostromedial tegmental nucleus (RMTg), dopaminergic neurons of the substantia nigra compacta (daSNC), and matrix neurons (mSPN). Numbers show the order of connection for the three subcircuits (1→2→3, 1→2→2B→3→4, 1→2C.1→2.C.2→2.C.3→3→4).

b, During decision-making, sSPNs help to determine which decision-dimensions are most important. Important decision-dimensions are later formed into decision-space (grids of cubes or squares) when dopamine is released from daSNC to mSPNs. For example, a 1D decision-space might be constructed where decision-making is conducted using information predominantly about reward, or multi-dimensional decision-space might be constructed for situations where multiple decision-dimensions are important to the decision.

c, After the decision-space is formed, mSPNs define action values (colors) based on cortical input (vectors) and the decision-space (grids). In a 1D decision-space, action values are assigned purely based on the activity of mSPN along a single decision-dimension. In multi-dimensional spaces, action values depend on the activity along multiple decision-dimensions.

d, GPi, LHb, RMTg, and daSNC activities affect the number of decision-dimensions that will be used to form the decision-space.

e,f, Circuit architecture (**e**) and functional role of circuit elements (**f**). The cortex encodes environmental and internal information. sSPN subpopulations each encode data along a decision-dimension. GPi→LHb→RMTg biases the circuit towards using more or fewer decision-dimensions. daSNC subpopulations, each corresponding to a decision-dimension, activate when their decision-dimension is important. This causes dopamine to be released to select mSPN subpopulations, and thus, a “decision-space” is formed from the basis of the important decision-dimensions. mSPNs define action values within the decision-space.

g, Scenario where the cortex encodes signals about food, social, and environmental cues. A “reward-predominant” decision-dimension captures information about music, cat, fire, and social interaction; a “cost-predominant” decision-dimension captures information about temperature and mosquitoes; and a “novelty-predominant” decision-dimension captures information about marshmallows. The most important decision-dimensions (here, the reward-predominant decision-dimension only, assigned the checkmark) are retained. mSPN forms action values using rules corresponding to the retained decision-dimensions, that is, within the decision-space. Several actions (here, “play music” and “stand up”) are assigned values, and then decisions are made based on the values of possible actions.

Figure 2

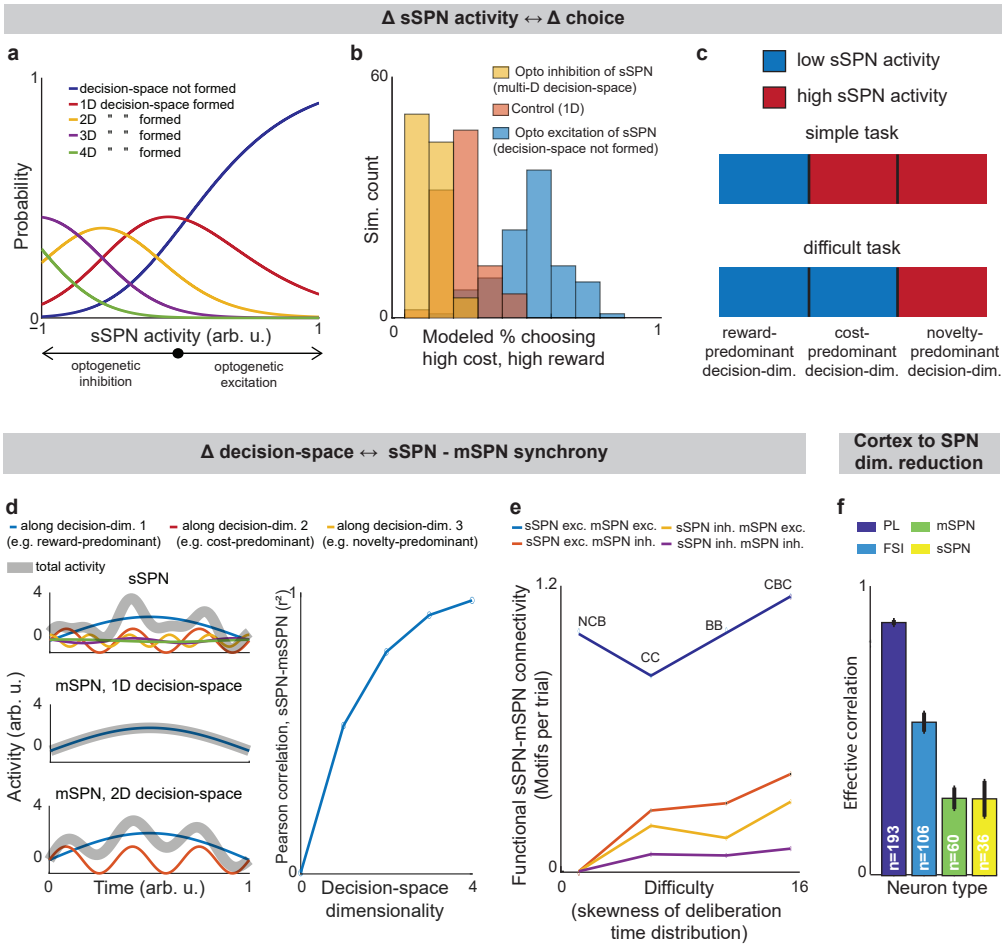


Fig. 2: Evidence of decision-space formation in sSPNs and mSPNs.

a, Changes to sSPN activity, for instance during optogenetic manipulation, cause changes in the modeled number of decision-dimensions used to form decision-space. In the analysis, b_{sSPN} in eq. (1) is incremented from -1 (low activity) to 1 (high activity).

b, More consistent choices are made at lower sSPN activity. 100 choices are simulated for each of the three decision-space scenarios. Modeled choices are made between a high-cost, high-reward option and a low-cost, low-reward option. The multi-D decision-space applies rules for both reward and cost, the 1D decision-space rules about only reward, and the "decision-space not formed" case applies neither. Then action values and choice are derived.

c, An sSPN subpopulation has reduced activity when it is used to form decision-space. Therefore, mean sSPN activity is different in simple tasks requiring low-dimensional decision-space versus difficult tasks requiring high-dimensional decision-space.

(legend continued on next page)

Fig. 2 continuation

d, When a one-dimensional decision-space is formed, sSPN (top-left panel) and mSPN (middle-left) activities have low correlation over time. As more decision-dimensions are used to form the decision-space, sSPN and mSPN activities increasingly correlate (bottom-left). Pearson's correlation between sSPN and mSPN activities (right panel).

e, Experimental data showing that task difficulty (measured through skewness of the deliberation time distribution) increases with functional sSPN-mSPN connectivity (measured through Granger causality) during a decision. Tasks: NCB = non-conflict cost-benefit (sSPNs = 14, mSPNs = 260), CC = cost-cost (sSPNs =46 , mSPNs = 400), BB = benefit-benefit (sSPNs =83 , mSPNs = 1246), CBC = cost-benefit conflict (sSPNs = 84, mSPNs =717). The CBC task has significantly more motif counts per trial ($p<0.003$, compared to shuffled data).

f, Experimental data showing that simultaneously recorded SPNs have less correlation (measured as effective correlation) than FSIs or prelimbic cortical neurons. This indicates dimensionality reduction from cortical neurons to SPNs. Significances of difference from cortical neurons: FSI $p<10^{-18}$, mSPN $p<10^{-45}$, sSPN $p<10^{-63}$.

Figure 3

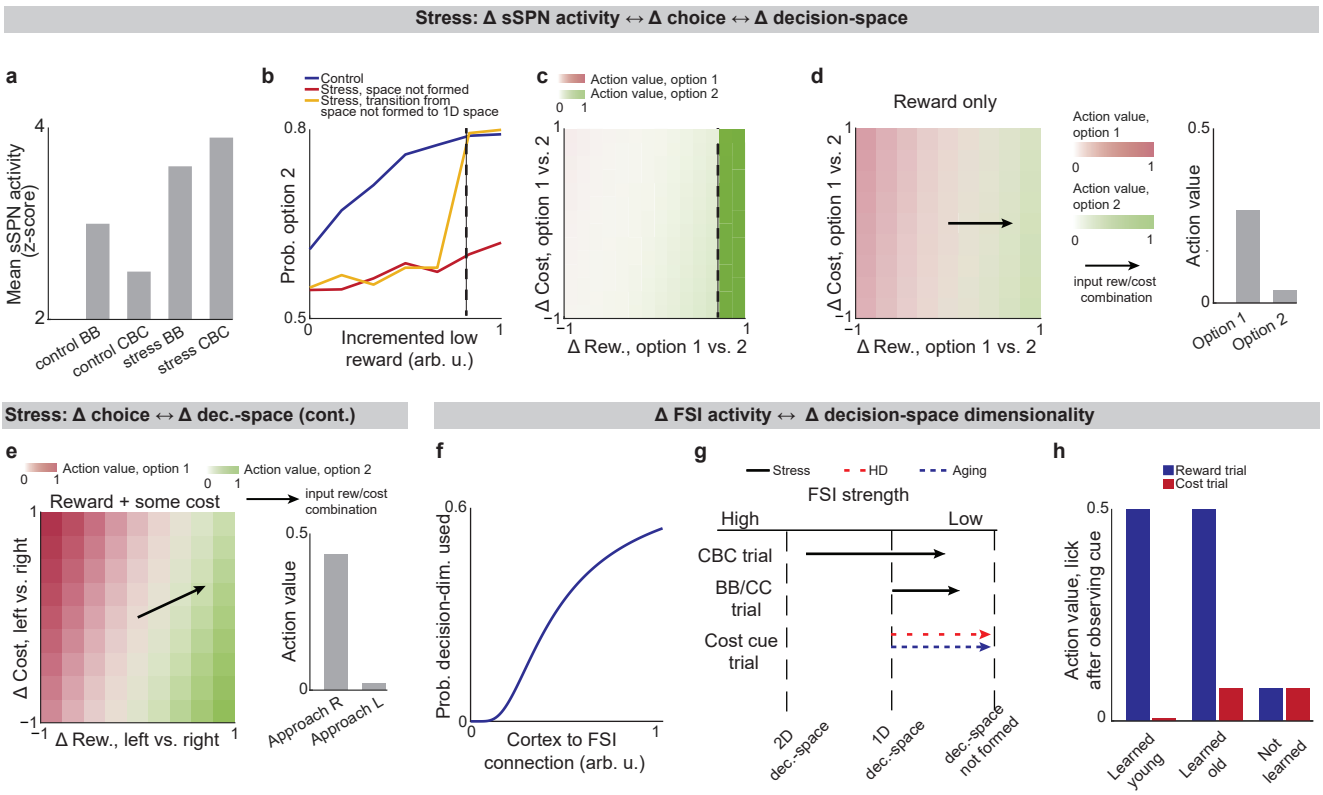


Fig. 3: Lower-dimensional decision-spaces are produced in stress, aging, and HD, affecting decisions.

Data was taken from experiments where rodents performed the cost-benefit conflict (CBC) task, in which rodents had to select between a high cost-high reward option and a lower cost-lower reward option. Two behavioral control tasks were also used: the benefit-benefit (BB) task, in which rodents selected between a high reward option and a low reward option (equal and minimal cost for both), and the cost-cost (CC) task, in which rodents selected between a high cost option and a low cost option (equal and minimal reward for both).

a, Summary of the experimental finding that sSPN activity is increased after stress, especially during the more difficult CBC task.

b,c, Modeled psychometric functions (**b**) and action values across experimental conditions (**c**) for a rodent in a T-maze task after chronic stress. The circuit forms a 1D decision-space only after reward exceeds a critical concentration (dashed line). At this point, the stress-group rodents switch from choosing the options roughly evenly to most often turning right towards the lower-cost option. Psychometric functions resemble experimental decision-making data.

d,e, After stress, a low-dimensional decision-space may be used by default (**d**) but a higher-dimensional decision-space may sometimes form during difficult tasks, for example those with both reward and cost (**e**). This leads to the counterintuitive result that adding cost to an offer (vectors on colormaps) can increase its action value (bar plots).

f, Cortex→FSI connection strength affects decision-space dimensionality by altering the propensity of a decision-dimension to form decision-space.

g, Cortex→FSI connection strength is reduced in stress, Huntington's disease, and aging. This leads to lower-dimensional decision-space.

h, Model of an operant conditioning task in Friedman et al (2020). Modeled action values for licking versus not licking in an operant conditioning task. There are two tasks: 1) responding to a reward cue by forming decision-space from a reward-predominant decision-dimension, and 2) likewise for cost. "Learned young" succeeds at 1 and 2, "learned old" at 1 but not 2, and "Huntington's disease" at neither. Resembles experimental licking rates.

Figure 4

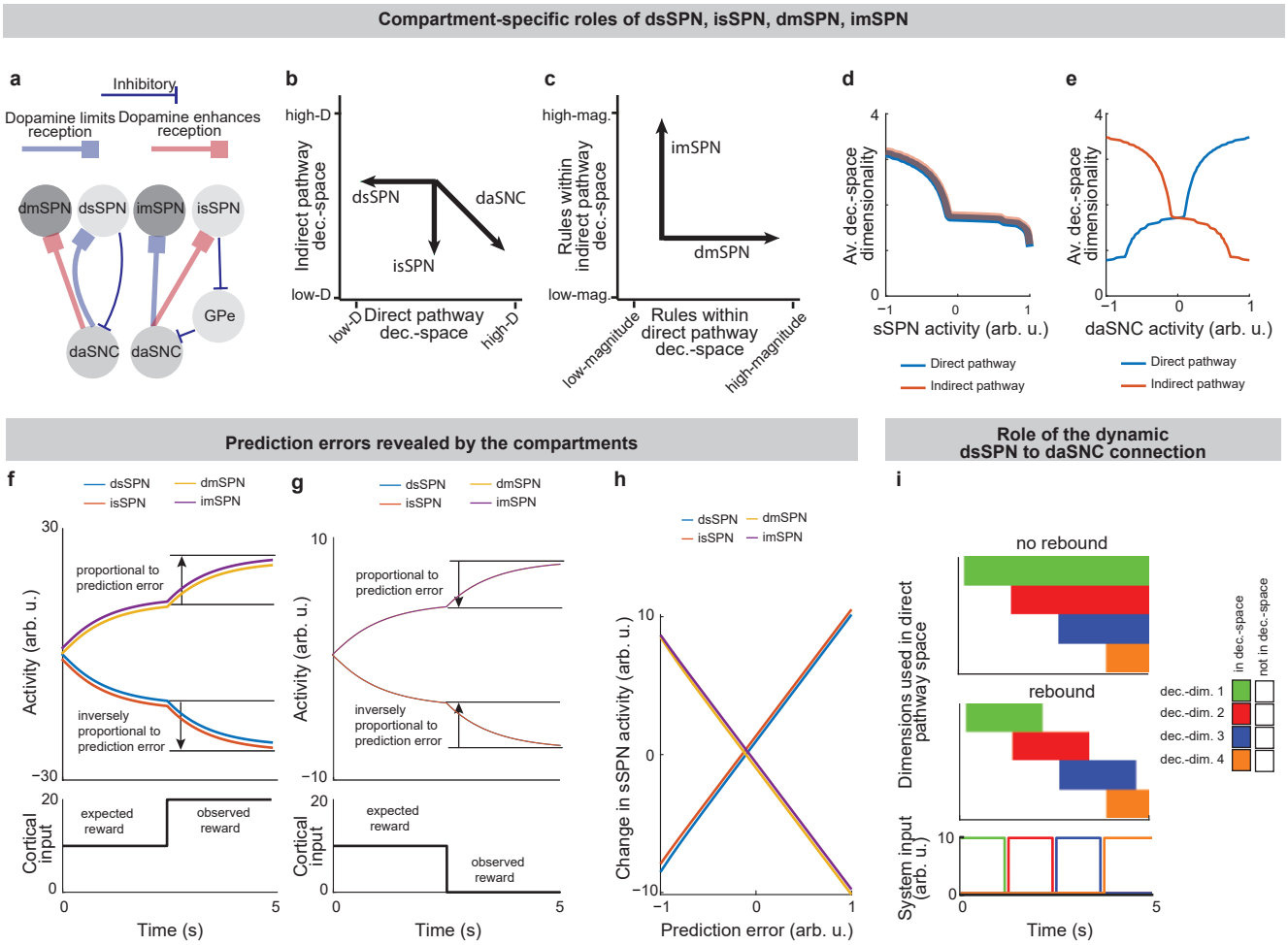


Fig. 4: dsSPNs, isSPNs, dmSPNs, and imSPNs serve unique roles in constructing and using the decision-space.

a, dsSPNs inhibit daSNC subpopulations and isSPNs disinhibit daSNC subpopulations via their connection through GPe. Dopamine lengthens upstates in dmSPN, enhancing reception of cortical signal, and shortens upstates in imSPN, limiting reception of cortical signal. For connectivity details, see **Table 2**.

b, As a result of the circuit connectivity, two decision-spaces are constructed in parallel, one in the direct pathway and one in the indirect pathway. Circuit elements, when active, bias the dimensionality of the direct pathway decision-space, the indirect pathway decision-space, or both (vectors).

c, Once the decision-spaces are formed, dmSPN activity amplifies the information along the direct-pathway decision-dimensions, and imSPN activity amplifies the information along the indirect-pathway decision-dimensions.

(legend continued on next page)

Fig. 4 continuation

d, Simulations where dsSPN or isSPN activity is incremented and the dimensionalities of the decision-spaces are counted at each time step. When overall sSPN activity changes, the dimensionalities of the direct and indirect decision-spaces change in the same direction. For example, at low sSPN activity, actions might be performed or refrained from based on careful consideration of many decision-dimensions, while at high sSPN activity, actions might be performed based on only reward or refrained from based on only cost.

e, Similar, but overall daSNC activity is incremented. When daSNC activity changes, the dimensionalities of the direct and indirect decision-spaces change in opposite directions. For example, at high daSNC activity, actions might be performed based on careful consideration of many decision-dimensions or refrained from based on only cost, potentially leading to impulsivity. Meanwhile, at low daSNC activity, actions might be performed based on only reward or refrained from based on careful consideration of many decision-dimensions, potentially leading to low motivation.

f,g, SPN signaling of positive (**f**) or negative (**g**) prediction error. Cortical input from one source, for instance a bell predicting a reward, begins at 0s. At 2.5s, reward information from a different source, for instance the administration of chocolate milk, is higher (**f**) or lower (**g**) than predicted by the first source, leading to a prediction error. In this example, the dsSPN, isSPN, dmSPN, and imSPN subpopulations corresponding to reward-related information reveal the prediction error, and not other SPNs.

h, The simulations in **g** and **h** are run for incremented differences in expected versus observed reward, from -1 (negative prediction error) to +1 (positive prediction error), demonstrating that each population changes its activity roughly proportionally to prediction error.

i, 1.25s pulses of input from the cortex (bottom panel) are applied along the four decision-dimensions in succession. A dynamic connection weight from dsSPN to daSNC facilitates the rapid de-prioritization of no longer required decision-dimensions in the “rebound” scenario, in which a period of dsSPN inhibition is followed by a rebound in daSNC activity above baseline levels (middle panel). This effect is removed in the “no rebound” scenario (top panel). In the model, the longer sSPN inhibits daSNC, the more the connection weight decreases in strength, leading to daSNC receiving decreasing signal about the importance of the decision-dimension. Then, when sSPN ultimately signals that the decision-dimension is no longer important, this signal is enhanced upon reception by daSNC.

Figure 5

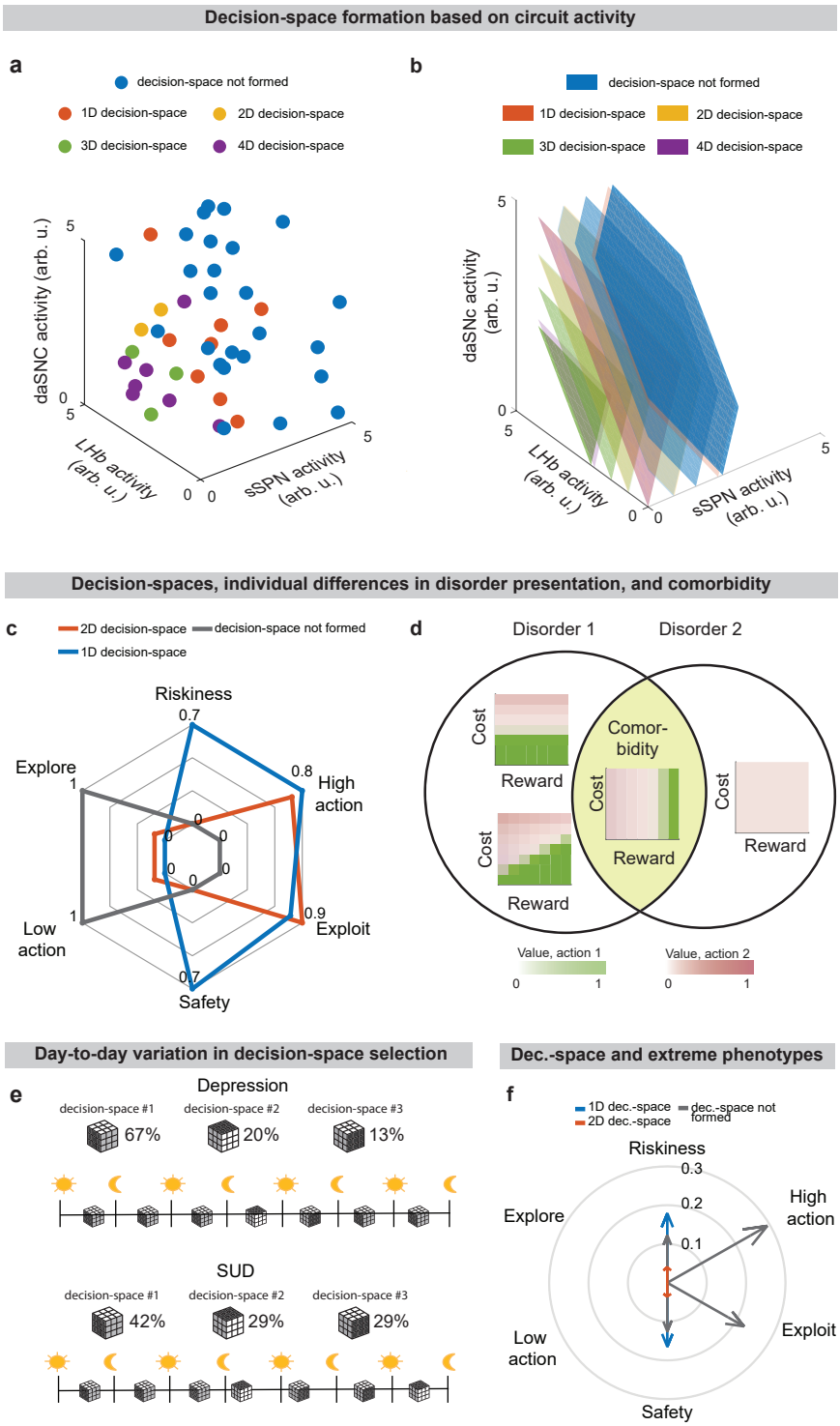


Fig. 5: Differences in decision-space could explain comorbidity, individual differences, and daily variations.

a, A decision-space is simulated for various sSPN, daSNC, and LHb activities. A high-dimensional decision-space (yellow, green, purple points) tends to form when sSPN activity is low, daSNC activity is high, and LHb activity is low, although there is an element of probability.

(legend continued on next page)

Fig. 5 continuation

b, The probabilities of the decision-spaces in **a** are shown using isosurfaces. Three isosurfaces are displayed for each decision-space dimensionality (probability = 0.25, 0.5, and 0.75 of formation, from lightest to darkest). The circuit forms a bias towards certain decision-spaces over others, but different decision-spaces can form at the same circuit activity.

c, Summaries of overall decision-making profiles across trials in a T-maze task, scored in six ways, showing that transitions between decision-spaces can lead to very different decisions. A disorder could produce a bias, for instance, towards low-dimensional decision-spaces, which would in turn alter the decision-making profile. Scores are formed by quantifying the trend of action values across reward and cost levels. Riskiness/safety: treatment of high-reward, high-cost (or low-cost, low-reward) levels; high/low action: tendency towards high (or low) action values; exploit/explore: tendency to focus on one action versus many (**see Extended Data Fig. 7c**).

d,e, Differences in circuit activity between individuals could lead to decision strategies observed at different rates, as is the case in individuals with disorder comorbidity **(d).** Further, day-to-day shifts in circuit activity shifts could cause stark differences in decision strategies between days (**e**). Cartoon shows the hypothetical use of different decision-spaces for depression and substance use disorders.

f, Certain decision-spaces more often lead to action values that are extreme (ratio formed over 1000 simulations, as scored using the metrics in **c**), a feature of disorders. Vector length corresponds to outlier rate (proportion of scores for each group that fall within the top 10% of observations across all groups).

Figure 6

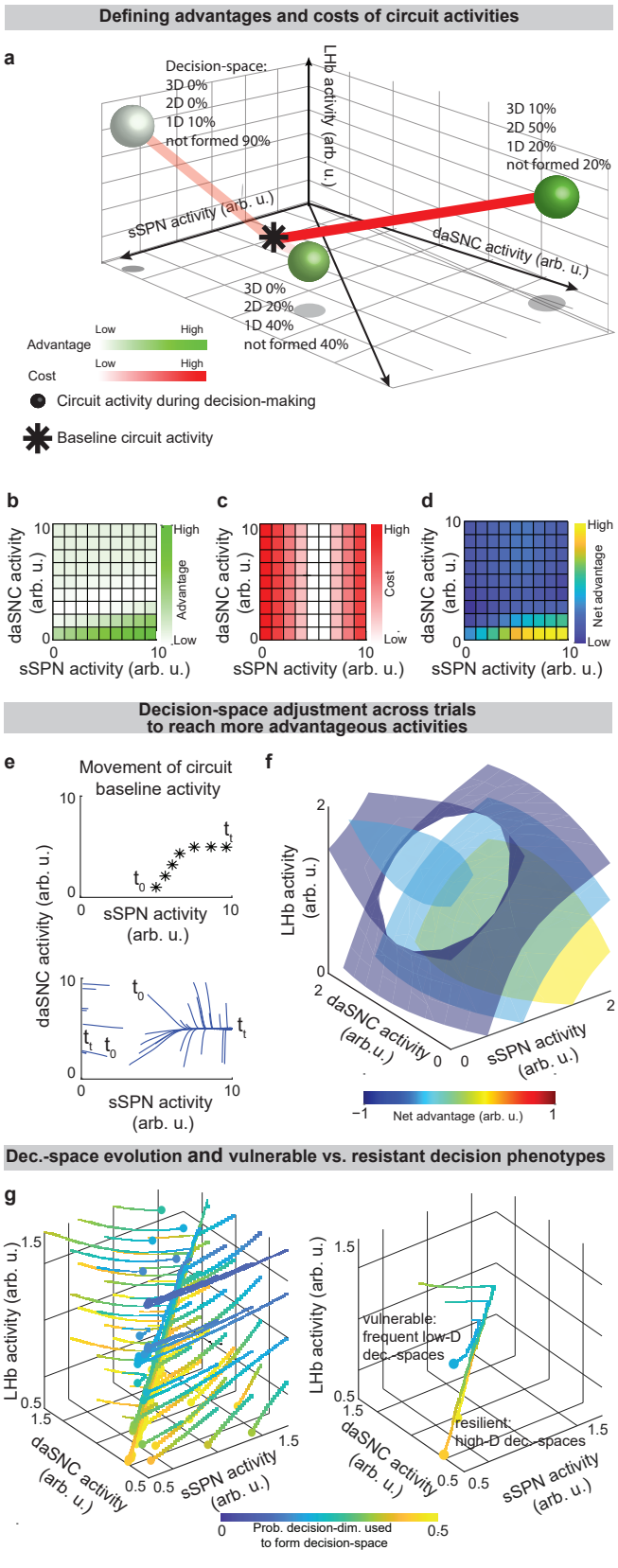


Figure 6 legend starts on next page

Fig. 6: Circuit adapts between trials to form preferred decision-spaces, leading to disorder progression.

a, Cartoon illustrating how the modeled circuit adjusts to facilitate construction of preferred decision-spaces ("advantage") while limiting large changes in circuit activity during decision-making ("cost"). Possible advantages and costs (whose difference is "net advantage") are shown at three circuit activities (balls). The decision-space is formed differently at each circuit activity (annotations), leading to differences in advantage. Between trials, the circuit adjusts its baseline activity (i.e. resting activity outside decision-making) in the direction of highest net advantage.

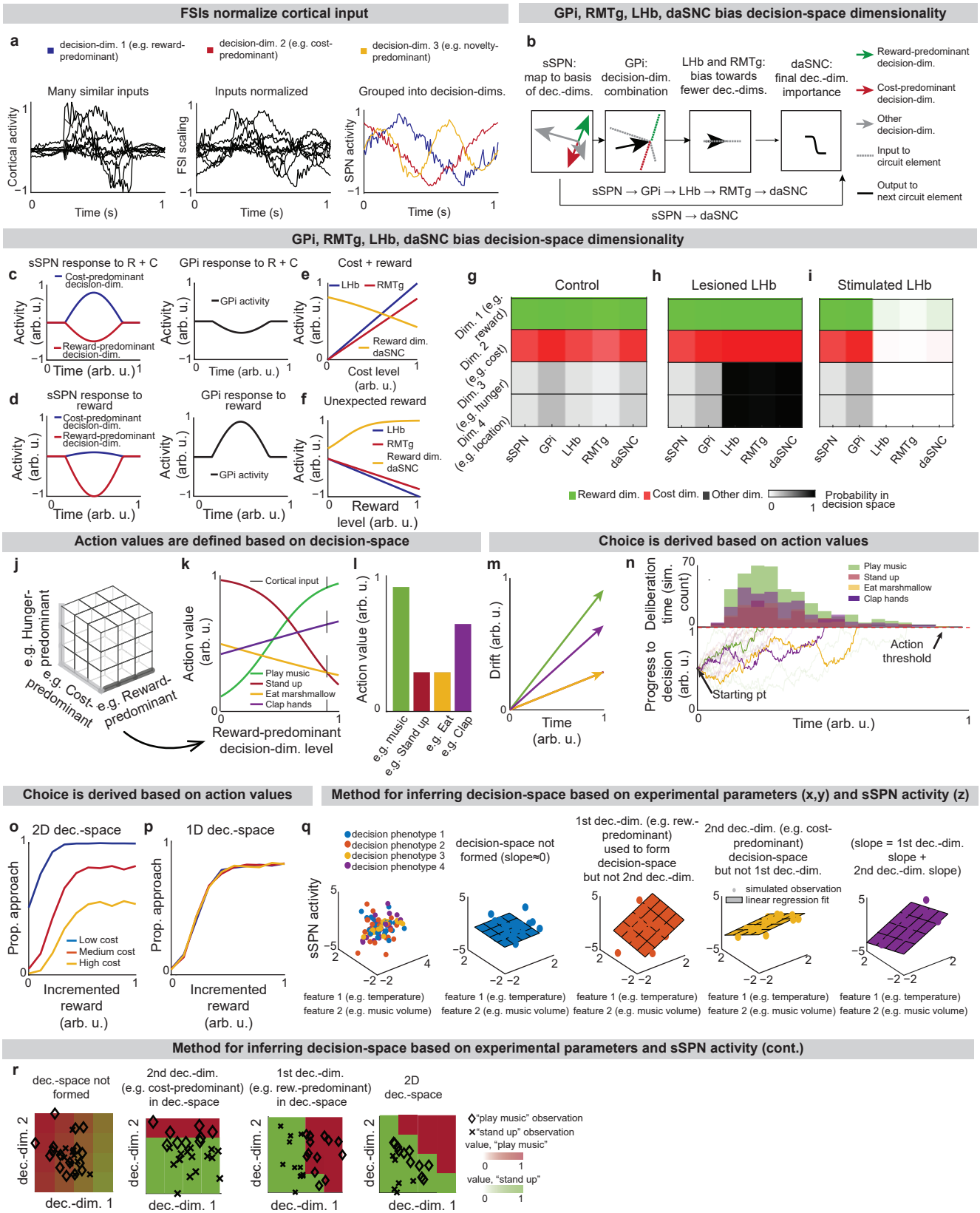
b-d, Simulated advantages (**b**), costs (**c**), and net advantages (**d**) shown across sSPN and daSNC activities.

e, sSPN and daSNC activity adjusts between trials to best form preferred decision-spaces. Due to this, the circuit adjusts its activity from an initial baseline activity (t_0) to other activities associated with the required decision-spaces (t_l).

f, Similar to **d** but for sSPN, LHb, and daSNC activities, showing a simulated circuit where it is most advantageous to have high sSPN activity during simple choices. The trend in the continuous 3D decision-space is visualized using isosurfaces.

g, Similar to **e** but for sSPN, LHb, and daSNC activities. Dots show ending circuit activities (i.e. t_l) and beginnings of lines possible initial circuit activities (i.e. t_0). Depending on initial activity, the circuit may increasingly use or disregard a decision-dimension when forming decision-space. Right panel shows trajectories from three starting circuit activities, two of which lead to the decision-dimension commonly being used to form decision-space (e.g. resilient subjects) and one which leads to it commonly being disregarded (e.g. vulnerable subjects).

Extended Data Fig. 1



Extended Data Fig. 1 legend starts on next page

Extended Data Fig. 1: Action values are defined within a decision-space formed by the circuit, Related to Fig. 1.

a, Example showing how FSIs and SPNs parse cortical activity in the model. The signals of 10 cortical neurons encoding sensory information (**a**) are normalized by FSIs such that activities are on a more uniform scale (**b**) and then mapped to sSPNs (**c**), which each encode activity along a "decision-dimension." In this figure (and **Fig. 1**), an instance of the model is described which allows for two convenient simplifications: 1) the representation of SPNs (both sSPNs and mSPNs) using one circuit element per decision-dimension, and 2) the treatment of the circuit using a feedforward model. See ***Instance 1: full connectivity and feedforward,*** **Methods**.

b, After FSI normalization, sSPNs map cortical activity to a basis of decision-dimensions. Mathematically, we represent, for a given pathway P (either direct pathway or indirect pathway), the mapping of cortical activity $\mathbf{x}_P$ to decision-dimensions as a linear transform via the matrix $\mathbf{W}_P$, whose columns correspond to the first several (in our analysis, 4) principal components of cortical activity. During this process, there is divisive normalization by FSI activity c_P and the potential for an overall shift in sSPN activity b_{sSPN} (eq. (1)). Next, GPi combines the sSPN signals into a single representation and RMTg and LHb bias this combined representation of signal along all decision-dimensions. We represent RMTg activity RMTg as the dot product of striosome to GPi weights $\mathbf{w}_{\mathrm{GPi}}$ and the activities the sSPNs $\mathbf{s}_{\mathrm{sSPN},P}$ for each pathway P, after incorporation of additive shifts z_{GPi}, z_{LHb}, and z_{RMTg} (eq. (5)). Next, daSNC takes input directly from sSPNs and from RMTg to calculate the final importance of each decision-dimension. We represent daSNC activity $\mathrm{daSNC}_{i,P}$ corresponding to decision-dimension i and pathway P as the activity of the corresponding sSPN population $\mathbf{s}_{\mathrm{sSPN},i,P}$ multiplied by a connection weight $w_{\mathrm{sSPN}\rightarrow\mathrm{daSNC},i,P}$, plus additive shifts applied individual to each daSNC element ($z_{\mathrm{daSNC},i,P}$) and uniformly to all daSNC elements (RMTg), all passed through an activation function (eq. (2)).

c,d, Example where sSPNs parse reward and cost (**c**), or reward (**d**) inputs along decision-dimensions (left) and then GPi combines the signals along the decision-dimensions into a single representation (right).

e,f, Modeled responses of LHb, RMTg, and daSNC to a cortical signal encoding reward and cost (**e**), and a cortical signal encoding an unexpected reward (**f**). LHb and RMTg increase their activities proportionally to the cost and decrease their activities proportionally to the reward. daSNC change their activities inversely. The modeled daSNC response is due to the combination of cortical inputs projected onto them directly from sSPN and through GPi, LHb, and RMTg.

g-l, Roles of the circuit elements in determining which dimensions form decision-space. Three scenarios are shown: control ($z_{\mathrm{LHb}}=0.5$ in eq. (5)), lesioned LHb ($z_{\mathrm{LHb}}=-5$), and stimulated LHb ($z_{\mathrm{LHb}}=5$). The colors shown for each circuit element correspond to the decision-space that would be formed absent the influence of downstream circuit elements.

j-l, Process by which action values are defined using decision-space. During a decision, a subset of decision-dimensions is selected, forming decision-space (**h**). This is represented mathematically through a diagonal matrix $\mathbf{S}_P$ whose elements are set probabilistically to either 1 (dimension in decision-space) or 0 (dimension not in decision-space) (eq. (3)). Otherwise, mSPN activity is formulated similarly to sSPN activity. Rules corresponding to retained decision-dimensions are used to define action values $v_{j,P}$ for action j and pathway P (**k,l**). This is represented mathematically as multiplication of a vector $\beta_{j,P}$ by mSPN activity $\mathbf{s}_{\mathrm{mSPN},P}$, subtracted by a shift $\alpha_{j,P}$ and run through an activation function (eq. (4)).

(legend continued on next page)

Extended Data Fig. 1 continuation

m,n, Action (or inaction) values $v_{j,P}$ for each action j and pathway P are used as drift rates (**m**) in a Merton process model (**n**). Discrete Merton processes obtained as a constant time step discretization of eq. (6) are run for each action simultaneously. At the time the first of the $v_{j,\text{direct}}$ processes reaches a defined threshold h, the corresponding action is enacted unless its inaction process $v_{j,\text{indirect}}$ has reached h first (eqs. (7),(8),(9)). Lines show the progress of processes towards a decision threshold for an example simulation. Histogram shows the decisions and deliberation times for the processes that reached the threshold first.

o,p, Psychometric functions derived across multiple reward levels. In the modeled experiment, the subject is asked to evaluate the rewards and costs of two offers and either approach or avoid. The decision-space (**o**: 1D, **p**: 2D) affects choice patterns.

q,r, Demonstration of a method by which decision-space can be inferred from sSPN activity, showing the utility of the mathematical formation of the model in connecting experimental inputs, sSPN activity, and choice. The method demonstrated here could be used to design an experiment in which the decision-space model is tested, or it could be used to explain differences in choice between groups, for instance control and disorder. In the simulation, two environmental inputs (e.g. temperature, music volume) across 100 sessions are classified based on decision-making phenotypes, for example based on reaction time, heart rate, eye movement (colors). Using the method visualized here, the decision-making phenotype classes are assigned one of four decision-space reference labels: 1) where decision-space is not formed, 2) where only the first decision-dimension is used to form decision-space, 3) where only the second decision-dimension is used to form decision-space, and 4) where both decision-dimensions are used to form decision-space. Synthetic data is generated by forming an arbitrary set of cortex→sSPN weights, using these to form sSPN activity, and then adding i.i.d. Gaussian noise. A linear regression is used to derive estimated decision-dimensions and assign hypothesized decision-spaces to each label. Choice is then examined with respect to the derived decision-dimensions. As expected, the regression slope (planes) corresponding to the 2D decision-space is roughly the sum of the regression slopes of the two 1D decision-spaces (**q**), and decisions correlate with the dimensions hypothesized to be used to form decision-space when choices are plotted against hypothesized dimensions (**r**). In the colormaps in **r**, action values are interpolated from an example set of observations (diamonds and Xs) via logistic regression, and a boundary line is drawn where the action value of “play music” equals the action value of “stand up.”

(1) $$\mathrm{s}_{\mathrm{sSPN},P} = \frac{1}{c_P}\mathbf{W}_P{}^{\mathrm{T}}\mathbf{x}_P + b_{\mathrm{sSPN}}$$ (copied for convenience)

(2) $$\mathrm{daSNC}_{i,P} = \frac{1}{1+\exp\left(w_{\mathrm{sSPN}\to\mathrm{daSNC},i,P}\cdot \mathrm{s}_{\mathbf{sSPN},i,P} + \mathrm{RMTg} - z_{\mathrm{daSNC},i,P}\right)}$$ (copied for convenience)

(3) $$\mathrm{s}_{\mathrm{mSPN},P} = \frac{1}{c_P}\,\mathbf{S}_P\mathbf{W}_P{}^{\mathrm{T}}\mathbf{x}_P$$ (copied for convenience)

(4) $$v_{j,P} = \frac{1}{1+\exp\left(-\beta_{j,P}\,\mathrm{s}_{\mathrm{mSPN},P} - \alpha_{j,P}\right)}$$ (copied for convenience)

(legend continued on next page)

Extended Data Fig. 1 continuation

(5) $$\mathrm{RMTg} = z_{\mathrm{RMTg}} + z_{\mathrm{LHb}} + z_{\mathrm{GPi}} \cdot \mathbf{w}_{\mathrm{GPi}} \cdot \begin{bmatrix} \mathbf{s}_{\mathrm{sSPN,\,direct}} \\ \mathbf{s}_{\mathrm{sSPN,\,indirect}} \end{bmatrix}$$

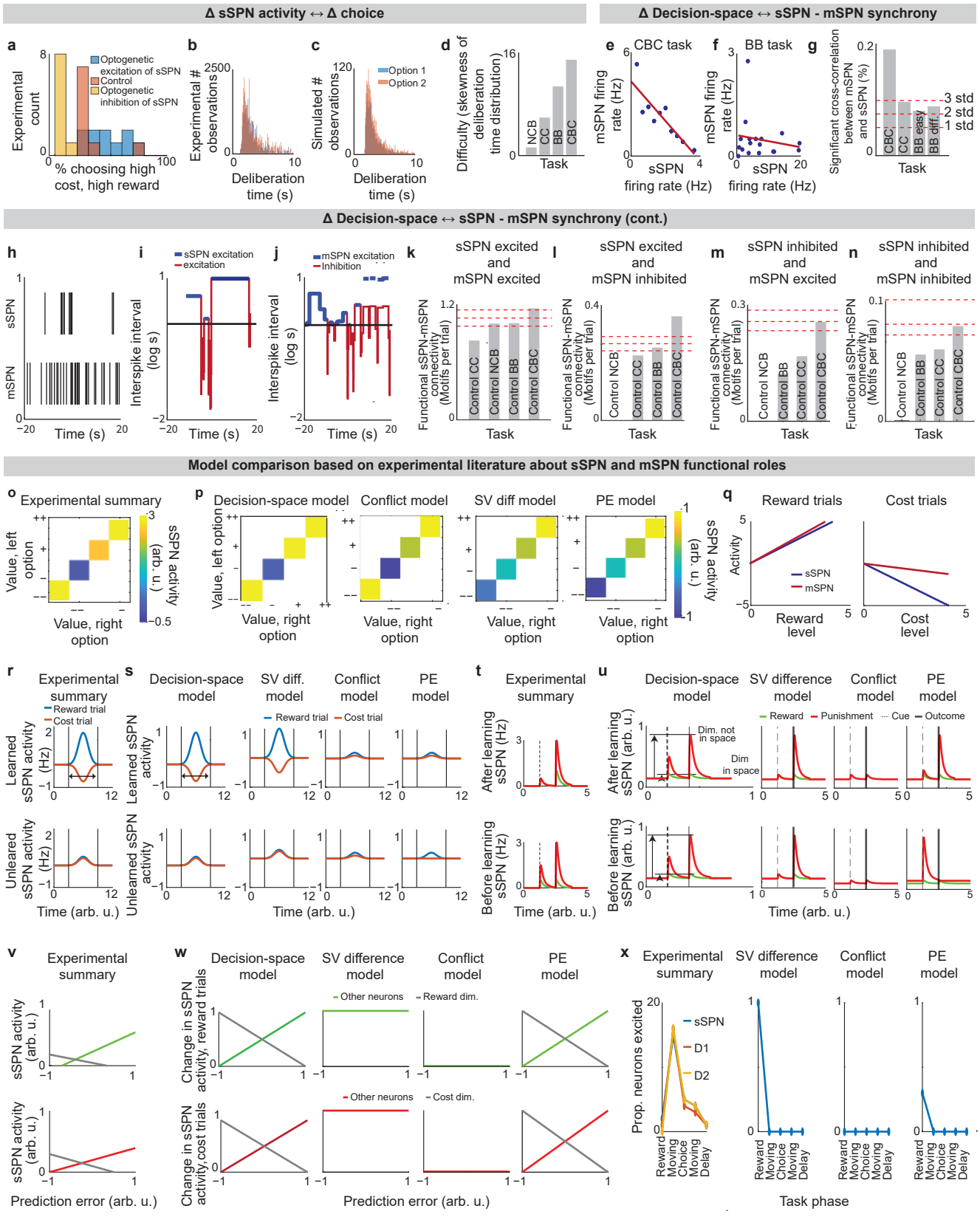
(6) $$dY_{j,P} = v_{j,P}\,dt + \sigma\,dW_{j,P},\ Y_{j,P}(t=0) = 0,\ \text{where } W_{j,P} \text{ is a standard Wiener process}$$

(7) $$t_{\mathrm{action},j} = \min_{t}\left\{ t \mid Y_{j,\mathrm{direct}} \geq h \right\}$$

(8) $$t_{\mathrm{inaction},j} = \min_{t}\left\{ t \mid Y_{j,\mathrm{indirect}} \geq h \right\}$$

(9) $$\mathrm{action} = \arg\min_{j \in J}\left(Y_j(t_{\mathrm{action},j}) \right),\ \text{where } J \text{ is the subset of actions s.t. } t_{\mathrm{action},j} < t_{\mathrm{inaction},j}$$

Extended Data Fig. 2



Extended Data Fig. 2 legend starts on next page

Extended Data Fig. 2: Decision-space model validation and comparison to alternative models, Related to Fig. 2.

a, Summary of experimental results of optogenetic manipulation during a conflict task in Friedman et al. (2015). Resembles the model in **Fig. 2b**.

b,c, Experimental deliberation time distribution (6 animals, 35 sessions) (**b**), which is successfully modeled using the Merton process model (**c**). Distribution is for the benefit-benefit task (concentration 70%). Experimental data here and throughout the figure is analyzed from the Corticostriosomal Circuit Stress Experiment database.

d, Skewness of the deliberation time distribution, which is used to estimate task difficulty. The CBC task had a deliberation time distribution that was more skewed than the non-conflict tasks. Tasks: NCB = non-conflict cost-benefit (4 rats, 27 sessions, 1250 trials), CC = cost-cost (7 rats, 25 sessions, 1852 trials), BB = benefit-benefit (7 rats, 128 sessions, 7762 trials), CBC = cost-benefit conflict (8 rats, 69 sessions, 3921 trials).

e,f, Analysis of relationship between sSPNs and mSPNs during decision-making. sSPN and mSPN neurons are significantly more correlated in tasks that require integration of reward and cost versus only reward or only cost. **e** shows representative examples from the cost-benefit conflict task (CBC, both reward and cost) and **f** shows the benefit-benefit task (BB, only reward), respectively.

g, The CBC task had significantly more correlated (Pearson's $r^2 > 0.4$ and significance < 0.05) pairs than the tasks which required integration of only reward or only cost. Confidence intervals (dashed red lines, 1,2,3 standard deviations) are estimated based on shuffled data. NCB = non-conflict cost-benefit (sSPNs = 14, mSPNs = 260), CC = cost-cost (sSPNs =46 , mSPNs = 400), CBC = cost-benefit conflict (sSPNs = 84, mSPNs =717), BB = benefit-benefit easy (sSPNs = 50, mSPNs =515 , chocolate milk concentration <50), BB = benefit-benefit difficult (sSPNs =33, mSPNs = 731, chocolate milk concentration >=50)

h-j, Process by which we identify functionally connected sSPN and mSPN neurons. Spiking times (**h**) are used to find inter-spike intervals (**i**,**j**) for sSPN (top rows) and mSPN (bottom rows) that were recorded simultaneously during decision-making. Intervals above the median interval (black line) are classified as inhibition (blue) and below median are classified as excitation (red). Functional connections are determined based on whether excitation or inhibition of sSPN followed by excitation or inhibition of mSPN.

k-n, Significantly more sSPN and mSPN neurons were functionally connected during decisions that required integration of reward and cost (CBC task) than the other tasks for all types of connections: sSPN excited and mSPN excited (**k**), sSPN excited and mSPN inhibited (**l**), sSPN inhibited and mSPN excited (**m**), and sSPN inhibited and mSPN inhibited (**n**). Tasks: NCB = non-conflict cost-benefit (sSPNs = 14, mSPNs = 260), CC = cost-cost (sSPNs =46, mSPNs = 400), BB = benefit-benefit (sSPNs =83, mSPNs = 1246), CBC = cost-benefit conflict (sSPNs = 84, mSPNs =717)

o, Mean sSPN activity should track decision-space dimensionality. Tasks in Friedman et al. (2015), plotted in order of difference between reward and cost, are assessed for likely decision-space dimensionality based on task difficulty (**Extended Data Fig. 2d**). Those with higher task difficulty are assigned lower sSPN activity per the model in **Fig. 2c**.

(legend continued on next page)

Extended Data Fig. 2 continuation

p, Three alternative models applied to the task in Friedman et al. (2015). The subjective value (SV) model assumes that sSPN encoded the relative action values of the two arms of the T-maze. The conflict model assumes that sSPN activity is inversely proportional to the amount of conflict in the task. The prediction error model compares expected value entering the task with reward or cost obtained on the maze. The conflict method resembles experimental results but not the SV difference model or the conflict model.

q,r, Summary of differences in sSPN activity across trials in the operant conditioning task in Friedman et al. (2020). mSPNs and sSPNs were active during reward trials, suggesting per the model in **Fig. 2c** that the decision-space was not formed (**q**). sSPNs were less active during the cost trials and were differently active than mSPNs, suggesting formation of a low-dimensional decision-space per the models in **Figs. 2c,d** (**r**)**.**

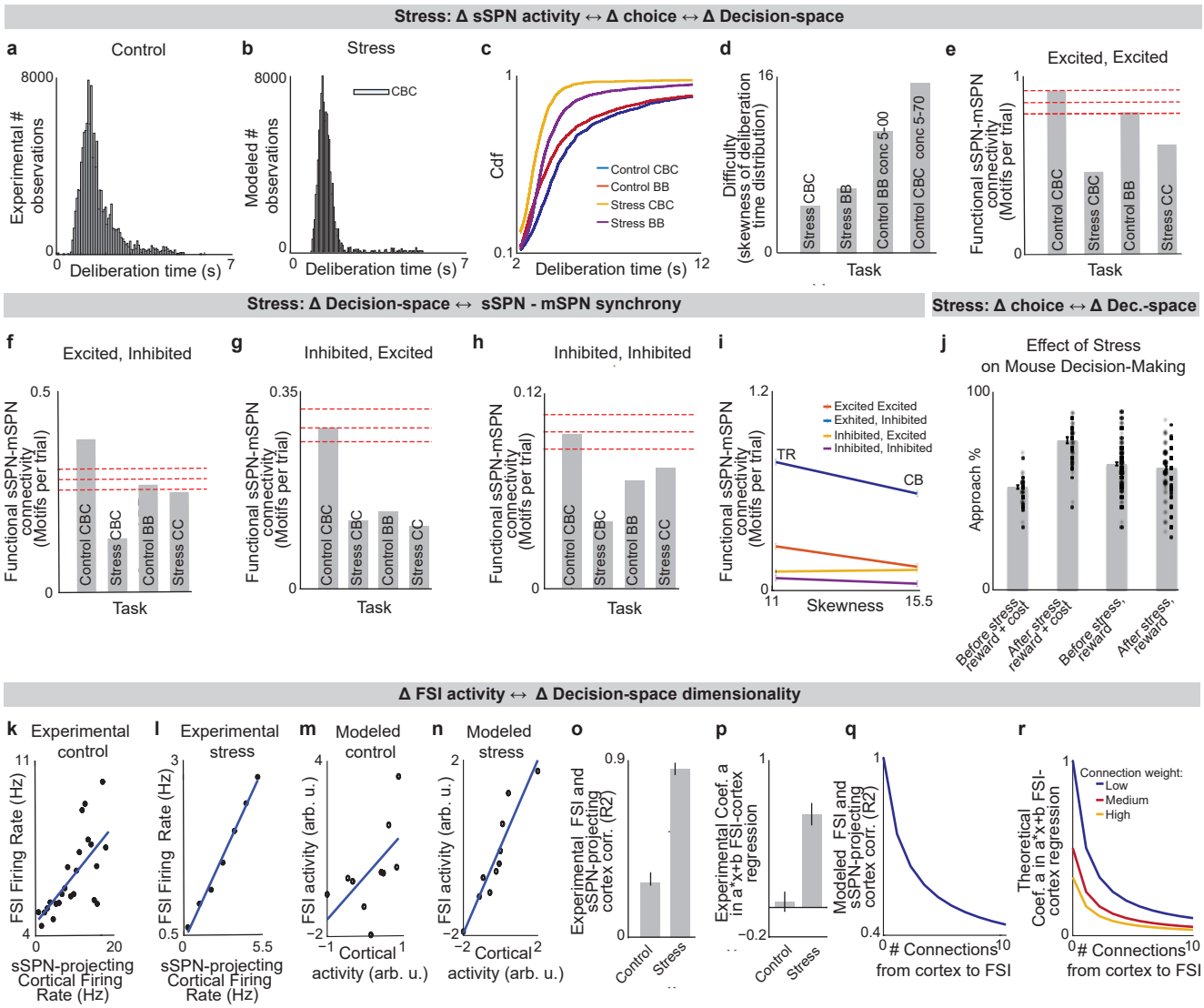
s, Prediction of sSPN activity for the task in **r** for the decision-space model and three alternative models. sSPN activity scales with overall subjective value in the task, so the subjective value model successfully interprets the experimental results.

t,u, Experimental results are summarized from Xiao et al. (2021) (**t**). Prediction of the decision-space model and three alternative models of sSPN subpopulation activity at the cue and during the outcome period of a Pavlovian conditioning task (**u**). Per the decision-space model, sSPN subpopulations should respond similarly to the cue and to the outcome because associated data is likely mapped along the same decision-dimension. The subjective value model expects more activity during the outcome period when reward is administered. The prediction error model expects more activity at the cue after learning. There was no conflict in the experimental setup.

v,w, Experimental results are summarized from Bloem et al. (2022) (**v**). Predictions of the decision-space model and the three alternative models of sSPN activity during a probabilistic Bandit task (**w**). The decision-space model expects that sSPN activity will reveal prediction errors, as does the prediction error model. The subjective value model instead anticipates that activity will track overall value regardless of prediction error. There was no conflict anticipated by conflict model.

x, In the value-guided choice task in Weglage et al. (2021), activities of all neuron types recorded (sSPN, dmSPN, imSPN) resembled one another over phases of the task and did not solely track subjective value, prediction error, or conflict. As shown in **Fig. 2d**, the experimental finding is an expectation of the decision-space model when a high-dimensional decision-space is formed.

Extended Data Fig. 3



Extended Data Fig. 3: Analysis of neural and decision-making data shows that decision-space is changed after stress, Related to Fig. 3. (84/85)

a,b, Deliberation time distributions of rodents performing the cost-benefit conflict task before stress (**a**) are less skewed and have shorter deliberation time than after stress (**b,** control: 8 rats, 198 sessions, 11683 cells; stress: 5 rats, 138 sessions).

c,d, Cumulative distribution functions (**c**) and distribution skewness (**d**) of deliberation time show that after chronic stress, the task involving integration of both reward and cost (CBC task) changes from producing the slowest (blue) to the quickest (yellow) decisions. CBC choice was slowest in control rats ($p < 0.0001$, KS-test). After stress, CBC choice was faster than in the other tasks ($p<0.0001$). Control CBC task: 69 sessions, 8 rodents; Control BB task: 128 sessions, 7 rodents; Stress CBC task: 34 sessions, 5 rodents; Stress BB task: 104 sessions, 5 rodents.

(legend continued on next page)

Extended Data Fig. 3 continuation

e-h, sSPNs and mSPNs were more functionally connected during a difficult (CBC) task before stress than after, per all possible types of connection: sSPN excited, mSPN excited (**e**), sSPN excited, mSPN inhibited (**f**), sSPN inhibited, mSPN excited (**g**), and sSPN inhibited, mSPN inhibited (**h**). Significance levels, depicted by dashed lines, show one (bottom), two (middle), and three (top) STD for functional connections calculated from shuffled control data. Control cost-benefit conflict (CBC) task: 92 pairs, 8 rodents; Tasks: Control BB = benefit-benefit (sSPNs =83, mSPNs = 1246), Control CBC = cost-benefit conflict (sSPNs = 84, mSPNs =717), Stress CBC = cost-benefit conflict (sSPNs =41, mSPNs =898), Stress BB = benefit-benefit (sSPNs = 156, mSPNs = 2813).

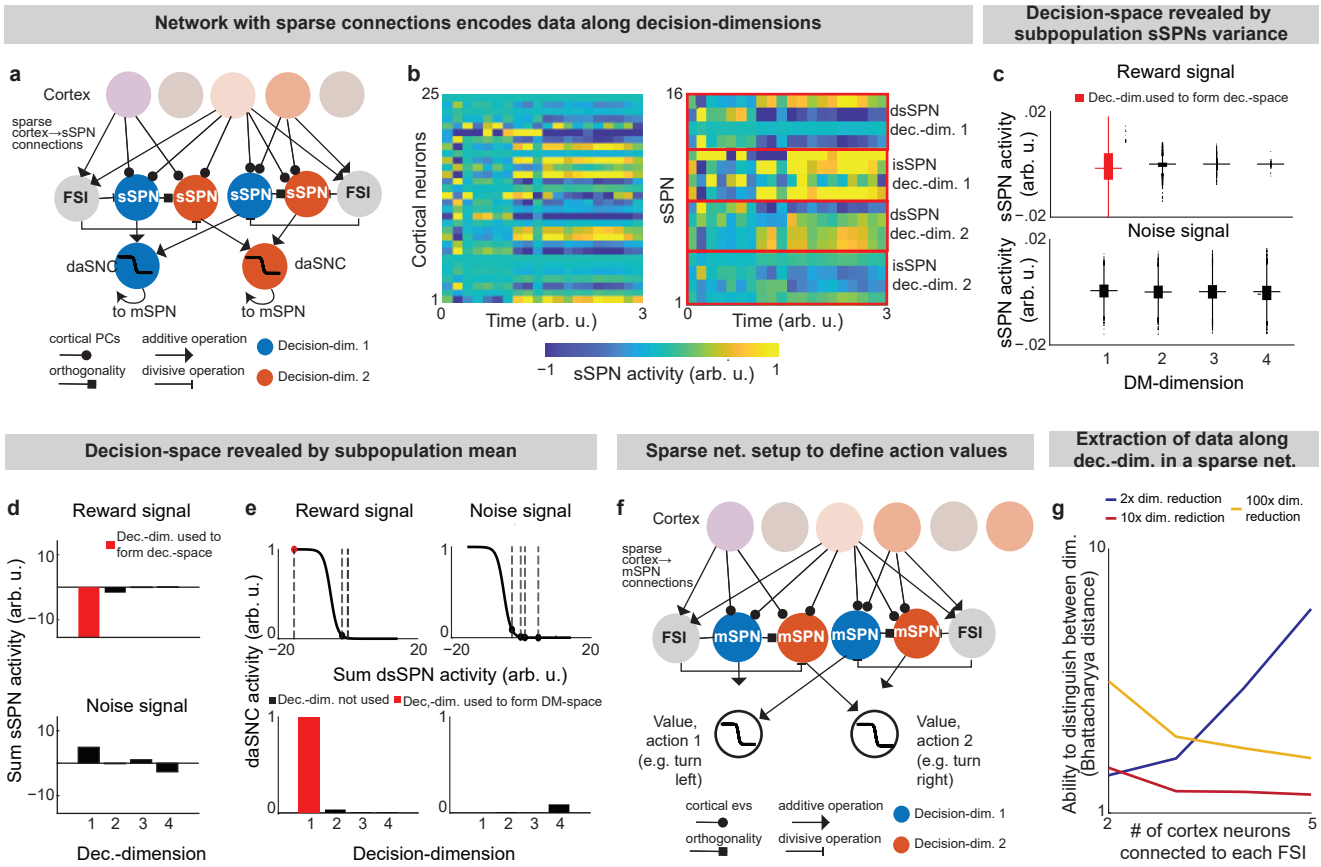
i, Deliberation time distribution skewness is no longer linked to functional connectivity after stress.

j, Experimental data that inspires the model in **Fig. 3h**. Mice that underwent chronic stress approached the lower-reward arm of the T-maze less when a small cost was added ("reward + cost" = cost-benefit conflict task, “reward” = benefit-benefit task). Dots = individual sessions, bar = mean across trials. There is significant difference ($p<10^{-19}$, paired t-test) in choice for the CBC task before and after stress, and nonsignificant difference in choice for the BB task before and after stress (p=0.10). CBC, before stress: 17 rodents, 38 sessions; CBC, after stress: 13 rodents, 34 sessions; BB, before stress: 23 rodents, 114 sessions; BB, after stress: 14 rodents, 116 sessions. Our model interprets this result as due to differences in the decision-space.

k-n, Representative examples of simultaneously recorded FSIs and prelimbic cortex neurons firing rates before (**k**, Pearson’s R=0.46) and after (**l**, R=0.99) chronic stress. After stress, there is less coordination between the connected pairs. This can be modeled as a reduction in the number of cortical neurons that synapse to each FSI (**m,n**).

o-r, In general, the neuron pairs in rats that underwent chronic stress had significantly higher correlation (**o**, $p<10^{-18}$) and significantly higher values of slope “a” in their $a \cdot x+b$ linear regression fits (**p**, $p<10^{-6}$). Control: 7 rodents, 78 neuron pairs; Stress: 4 rodents, 37 neuron pairs. This suggests, per our modeling, that there are fewer connections from cortical neurons to FSI after stress. Modeled squared Pearson correlation coefficient (**q**) and slope “a” parameter in the $a \cdot x+b$ fits (**r**) are shown when the connection between cortical and FSI neurons are altered in two ways: 1) through a reduction in the number of connections, and 2) through a reduction in the strength of each connection (i.e. connection weight). This experimental evidence is aligned with a reduction in the number of connections, suggesting that FSI normalization is disrupted after stress, leading to higher sSPN activity and thus formation of lower-dimensional decision-spaces.

Extended Data Fig. 4



Extended Data Fig. 4: Mean and variance of SPN activities reveals decision-space.

a,b, In this figure, we demonstrate that SPNs can encode activity along decision-dimensions successfully even in large, sparse networks, as exist in the brain, by considering an instance of the model where there are sparse connections between cortical neurons (n=50), FSIs (n=10,000), and SPNs (40,000 sSPNs, 40,000 mSPNs). sSPNs each encode activity along a principal component of a randomly sampled set of cortical neurons C (**a**). So, when activity changes in C (for example, during the reward cue in **b,** left panel), sSPNs (**b,** right panel) that each encode data along an ith principal component respond somewhat similarly to one another. See ***Instance 2: sparse connectivity and feedforward*, Methods**.

c, Decision-space is revealed from the variance of sSPN activities, aligning to the experimental result in **Fig. 2e**. Modeled sSPN subpopulations have greater variance when there is high-magnitude cortical signal along a corresponding decision-dimension (e.g., an sSPN subpopulation corresponding to a reward-predominant decision-dimension in response to a reward cue). Summaries are shown of the activities of 10,000 dsSPN (top row). Two types of cortical signal are passed to SPNs: one with reward (left column, matches the example used in **Fig. 4**) and one with Gaussian white noise (right column). To produce the network, 10,000 groups of 4 randomly sampled cortical neurons (notated as the set C) are connected to an FSI and 4 dsSPNs (or isSPNs) and 4 dmSPNs (or imSPNs) for each pathway. The activity on a given sSPN which receives projection from C, $\mathrm{sSPN}_{s,C}$, is defined mathematically based on the weights from cortical neuron q to sSPN s, $w_{q\to s}$, the activity of the FSI which received projection from C, FSI_C, and an additive shift that represents the relative activity of all sSPN neurons, b_{sSPN} (eq. (10)).

(legend continued on next page)

Extended Data Fig. 4 continuation

d, Decision-space is revealed by the mean of sSPN activities, aligned with the model in **Fig. 2c**. An sSPN subpopulation has lower mean activity when there is high-magnitude cortical signal along a corresponding decision-dimension.

e, Selection process by which daSNC neurons determine which decision-dimensions to include and to not include in decision-space, illustrating the construction of decision-space from sSPN activity. Lines are daSNC activation functions. Placement along the x-axis (emphasized by dashed lines) is the subpopulation average activity in **d**. Mathematically, daSNC neuron corresponding to decision-dimension i and pathway P, $\mathrm{daSNC}_{i,P}$, is computed as the weighted average of sSPNs corresponding to that decision-dimension and pathway, shifted by RMTg activity RMTg and a daSNC biasing factor $z_{\mathrm{daSNC},i,P}$, all passed through an activation function (eq. (11)).

f, Correlate to the sSPN-centered subnetwork described in **a** for mSPN, illustrating how action values could be defined by a network with sparse connections. The activity of a given mSPN which receives projection from C, $\mathrm{mSPN}_{m,C}$, is defined similarly to an sSPN but for term representing dopamine signaling from the daSNC neuron corresponding to decision-dimension i and pathway P to an mSPN corresponding to the same decision-dimension and pathway, $d_{i,P}$ (eq. (12)).

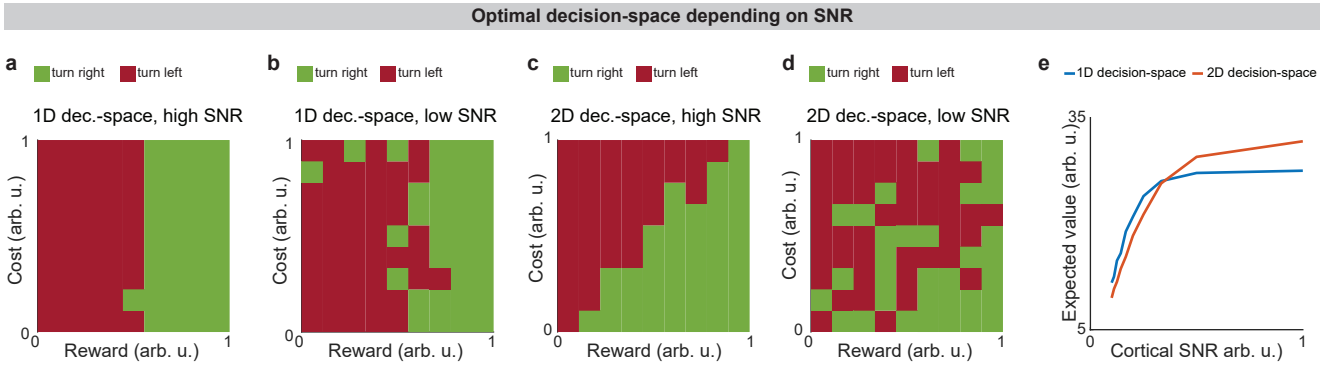
g, Bhattacharya distance between the distributions of SPNs encoding data along the reward-predominant and cost-predominant decision-dimensions. sSPN can correctly differentiate reward from cost signal despite sparse cortico-striatal connectivity. Lines show the averages of 1000 simulations. This result demonstrates the feasibility of data along decision-dimensions being encoded by neural populations with sparse connections.

(10) $$\mathrm{sSPN}_{s,C} = \frac{1}{|C|}\sum_{q\in C}\frac{w_{q\to s}\,\mathrm{cortex}_q}{\mathrm{FSI}_C} + b_{\mathrm{sSPN}}$$

(11) $$\mathrm{daSNC}_{i,P} = \frac{1}{1+\exp\left(\frac{1}{n_{\mathrm{sSPN}}}\sum_{s\in i,P} w_{s\to \mathrm{daSNC},i,P}\cdot \mathrm{sSPN}_s + \mathrm{RMTg} - z_{\mathrm{daSNC},i,P}\right)}$$

(12) $$\mathrm{mSPN}_{m,C} = \frac{d_{i,P}}{|C|}\sum_{q\in C}\frac{w_{q\to m}\,\mathrm{cortex}_q}{\mathrm{FSI}_C}$$

Extended Data Fig. 5

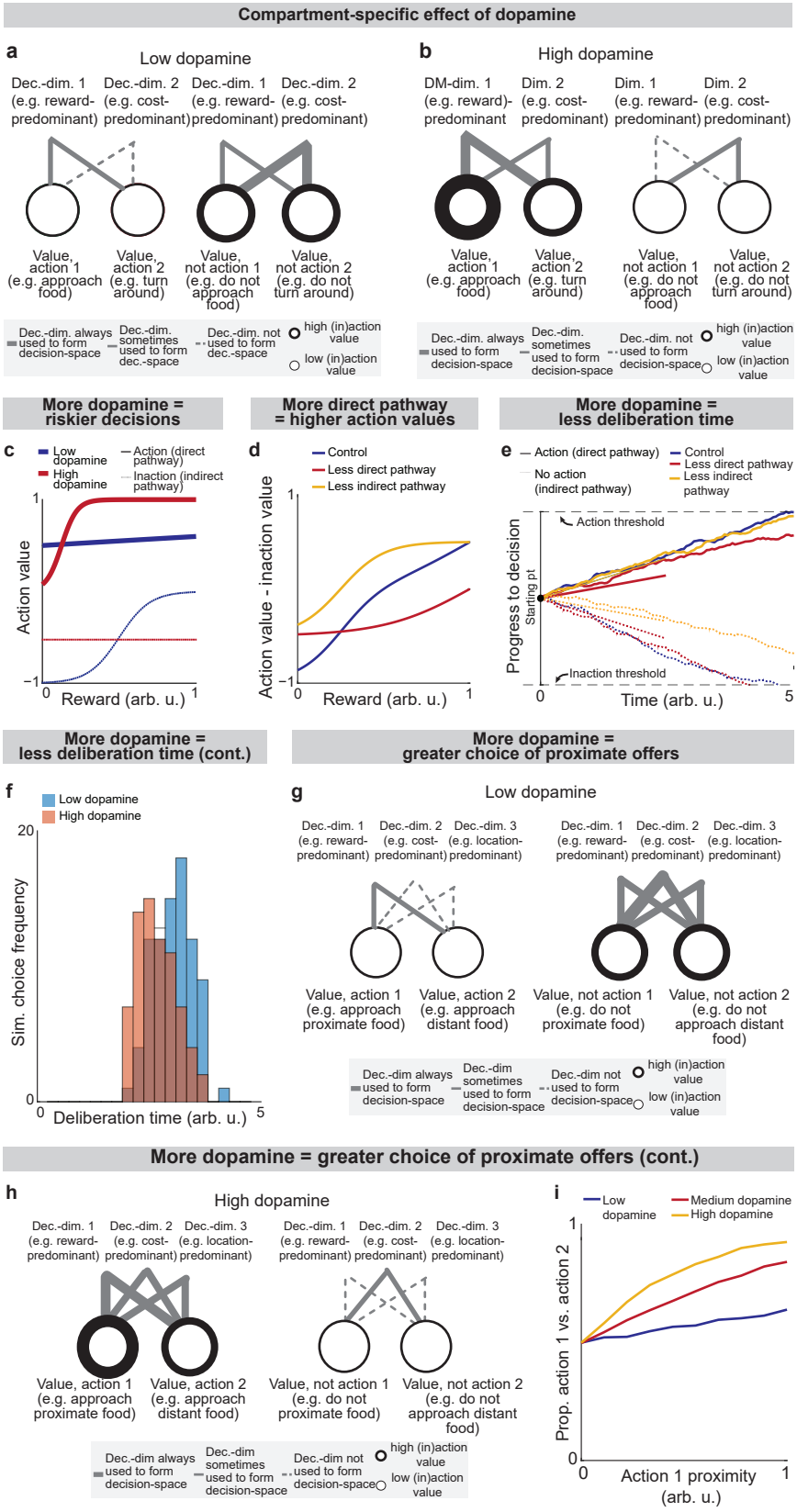


Extended Data Fig. 5: Decision-space is differentially constructed based on cortical signal-to-noise ratio (SNR).

a-d, Modeled T-maze task where an animal turns right to choose a reward/cost offer or turns left to avoid it. Cases where there is high cortical SNR (**a**,**c**) or low cortical SNR (**b,d**) and a 1D decision-space (**a,b**) or 2D decision-space (**c,d**) are formed. Modeled "turn right" actions are considered successful (positive value) when reward > cost, and "turn left" when reward < cost. 2D decision-spaces lead to more value when there is low cortical SNR (**c**) but not when there is high SNR (**d**).

e. Choices using different types of decision-spaces have different expected reward minus cost (expected value), depending on cortical SNR.

Extended Data Fig. 6



Extended Data Fig. 6 legend starts on next page

Extended Data Fig. 6: Roles of the direct and indirect pathways, Related to Fig. 4.
Here, analysis is conducted using the instance of the model described in **Extended Data Fig. 1.**

a,b, Effect of low (**a**) versus high (**b**) dopamine release on decision-space formed by the direct pathway (left panels) or indirect pathway (right panels). When dopamine release is low (**a**), low-dimensional direct pathway decision-spaces are constructed by dsSPNs and high-dimensional indirect pathway decision-spaces are constructed by isSPNs. The opposite happens when there is high dopamine (**b**). To analyze these effects, analysis in **Fig. 5** uses an instance of the model where the circuit elements interact dynamically, represented mathematically through a system of differential equations, where sSPN activity for a given decision-dimension i and pathway P, $s_{i,P}(t)$, respond to cortical input after normalization by FSI, $x_{i,P}(t)$, based on weights w between sSPNs, mSPNs, and daSNC elements, a decay factor τ, and a coefficient that controls sSPN→daSNC plasticity, κ (eqs. (13),(14),(15),(16)). The weight of a decision-dimension in mSPN, $S_{i,P}(t)$, occurs dynamically depending on whether or not dopamine release is above a specified threshold (eq. (17)). See ***Instance 3: full connectivity and dynamics,*** **Methods**. In the current figure, we conduct computational analysis using ***Instance 3****,* as defined in **Extended Data Fig. 1**.

c, Modeled action values for the cost-benefit conflict task where an incremented reward (from a low reward of 0 to a large reward of 1 arb. u.) is accompanied by a constant cost (set to 0.25 arb. u.). Increased dopamine leads to increased action value of high-reward, high-cost options and decreased “inaction value,” thus leading to more approaches when there is high reward and high cost.

d, The effects of the direct versus indirect pathways are examined in the cost-benefit conflict task used in **Figs. 2 and 3**. Modeled action values change when dmSPN and dsSPN or imSPN and isSPN are inactivated during a task with incremented reward (from a low reward of 0 to a large reward of 1 arb. u.) with medium cost (set to 0.5 arb. u.). Inactivated dmSPNs and dsSPNs lead to lower action values and inactivated imSPNs and isSPNs lead to larger action values.

e,f, Modeled deliberation time for the cost-benefit conflict task is shorter when there is more dopamine. Choice is modeled from action values calculated in the cost-benefit conflict task (here, reward = 1 arb. u., cost = 0.5 arb. u.). The method to derive choice from action values (upward-sloping drifts) and inaction values (downward-sloping drifts) is shown in **e** and modeled deliberation times are shown in **f**.

g-i, The model predicts that dopamine biases actions that contain rewards in physical and/or conceptual proximity. Decision-dimensions important for some decisions but not all will be used to derive action values when there is high dopamine (**g**) but not when there is low dopamine (**h**). If these decision-dimensions correspond to information about location, for instance, then additional dopamine may lead to the incorporation of spatial information in decisions, leading to actions containing the same location information having more similar action values (**i**).

(legend continued on next page)

Extended Data Fig. 6 continuation

(13) $$\tau \cdot \frac{\mathrm{d}s_{\mathrm{sSPN},i,P}(t)}{\mathrm{d}t} = -s_{\mathrm{sSPN},i,P}(t) - x_{i,P}(t) - w_{\mathrm{daSNC}\to\mathrm{sSPN},i,P} \cdot \left(y_{\mathrm{sSPN},i,P}(t) - \frac{1}{2}\right)$$

(14) $$\tau \cdot \frac{\mathrm{d}s_{\mathrm{mSPN},i,P}(t)}{\mathrm{d}t} = -s_{\mathrm{mSPN},i,P}(t) + x_{i,P}(t) + w_{\mathrm{daSNC}\to\mathrm{mSPN},i,P} \cdot \left(y_{\mathrm{sSPN},i,P}(t) - \frac{1}{2}\right)$$

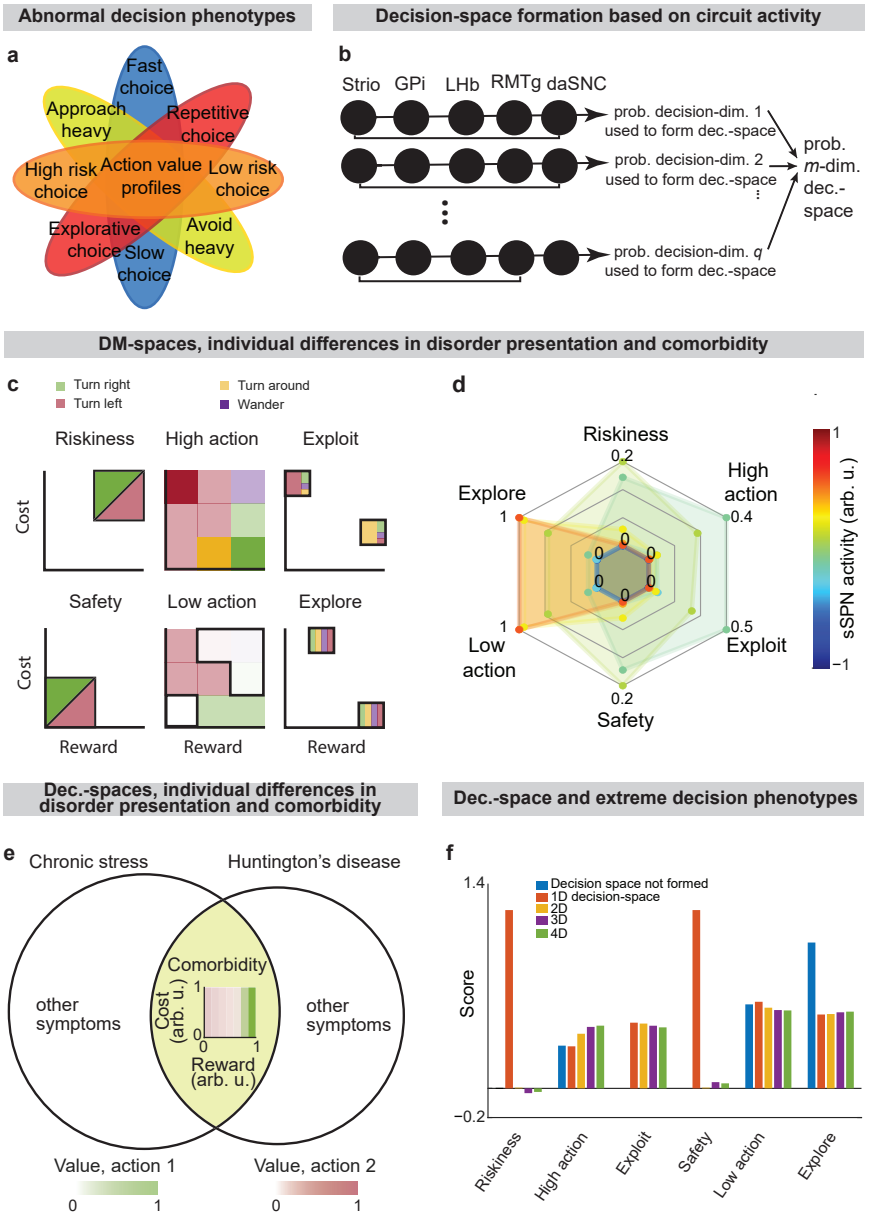
(15) $$\frac{\mathrm{d}}{\mathrm{d}t} w_{\mathrm{sSPN}\to\mathrm{daSNC},i,P}(t) = \kappa \cdot s_{\mathrm{sSPN},i,P}(t)$$

where:

(16) $$y_{\mathrm{sSPN},i,P}(t) = \frac{1}{1+\exp\left(w_{\mathrm{sSPN}\to\mathrm{daSNC},i,P}(t) \cdot s_{\mathrm{sSPN},i,P}(t) + \mathrm{RMTg} - z_{\mathrm{daSNC},i,P}\right)}$$

(17) $$\mathrm{S}_{i,P}(t) = \begin{cases} 0 & y_{i,P}(t) < \text{threshold} \\ 1 & y_{i,P}(t) \geq \text{threshold} \end{cases}$$

Extended Data Fig. 7



Extended Data Fig. 7: Individual decision-making differences can be explained by differences in decision-space, Related to Fig. 5.

a, Decision-making symptoms observed in disorders.

b, Schematic of the computational model for simulating decision-space based on circuit activity in the analyses plotted in **Figs. 7a,b**. daSNC activity $\mathrm{daSNC}_1 = \mathrm{daSNC}_2 = \ldots = \mathrm{daSPN}_q = d$ is calculated for each of q elements, each corresponding to a decision-dimension. Decision-space dimensionality is determined from the Poisson binomial distribution of individual decision-dimension probabilities, per eq. (18). See **Defining mSPN activity and decision-space, Methods.**

(legend continued on next page)

Extended Data Fig. 7 continuation

c, Illustration of the method by which we score subjective valuations along six axes (riskiness, safety, high action, low action, exploit, explore) for a modeled T-maze task. A grid of action values for a range of reward and cost combinations is formed. The “riskiness” and “safety” axes are calculated based on the action values for the high-reward high-cost combinations and low-reward low-cost combinations, respectively. “High action” and “low action” axes are calculated as the proportion of the reward/cost grid with especially high (sum of all action values > 0.5) or low (< 0.2) action value, respectively. The “exploit” and “explore” axes are determined as the proportion of the reward/cost grid with especially high (> 0.5) or low (< 0.25) Gini coefficients between the action values.

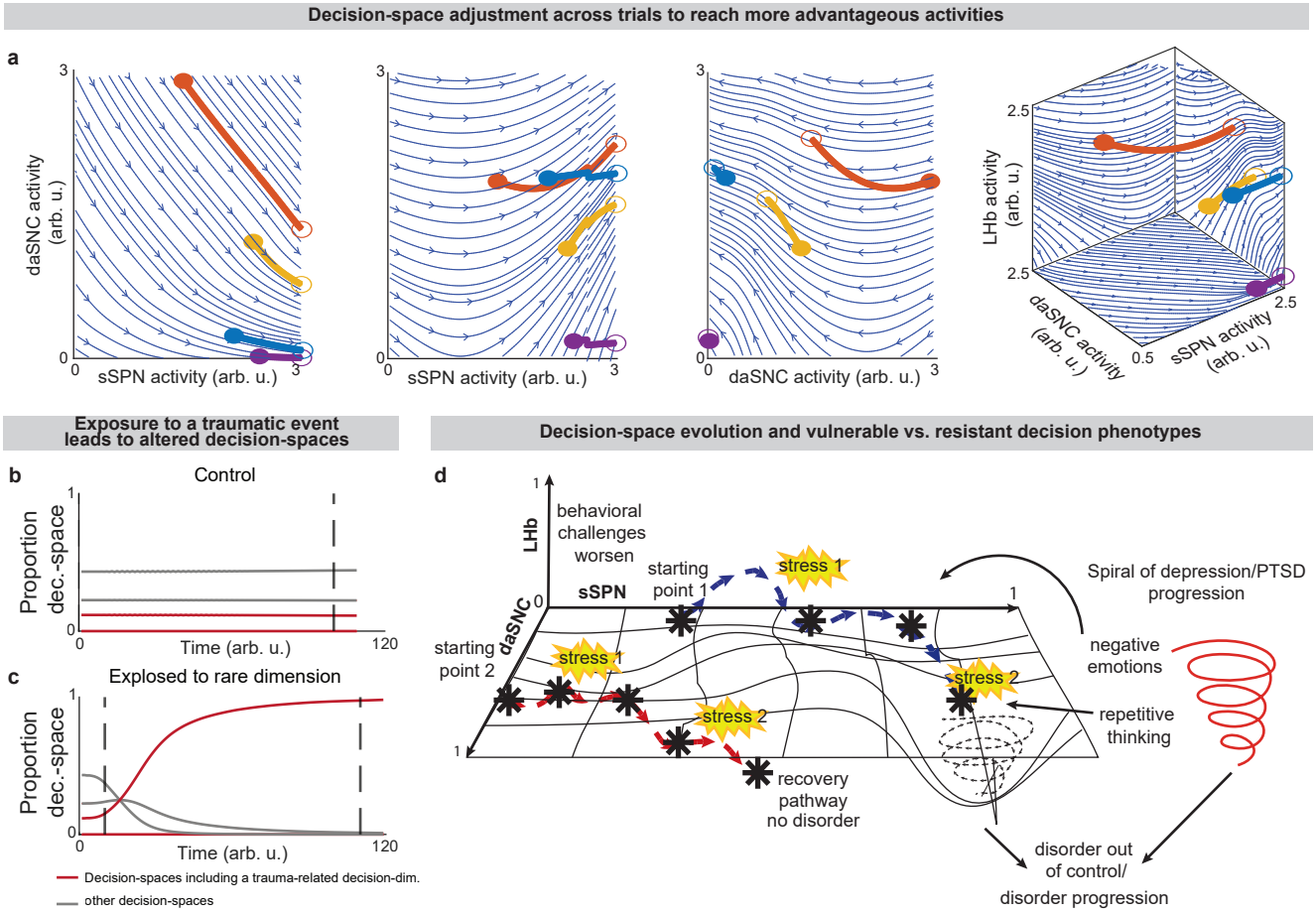
d, Extension of **Fig. 7c**. sSPN activity is modified, leading to decision-spaces formed at different rates. These different decision-spaces lead to different action valuations. Thus, in a disorder which affects sSPN activity during decision-making, differences in decision-space may be responsible for differences in decision-making.

e, Huntington’s disease and chronic stress both have decision-making signatures of low-dimensional decision-spaces.

f, Extension of **Fig. 7f**, showing that scores along the six subjective valuation axes (mean taken over 1000 simulations) are different depending on decision-space. Thus, disorders that affect decision-space formation may lead to shifts in decision-making.

(18) $$\mathrm{P}(\,m \text{ DM-dimensions used to form DM-space}) = \binom{q}{m} d^{m}(\,1-d)^{\,q-m} \quad \text{for } m = 0, 1, \ldots, q$$

Extended Data Fig. 8



Extended Data Fig. 8: Disorder development and decision-space formation, Related to Fig. 6.

a, Simulations of trajectories of circuit activity between trials. Four trajectories of circuit activity movement (thick lines) are plotted above streamlines (blue lines) for a modeled circuit that adjusts to facilitate simple choices. Filled circles are various initial circuit activities (t_0). Empty circles are the ending circuit activities (i.e. t_t). The left three panels are two-dimensional slices of the plot on the right. Here and in **Fig. 6**, the modeled circuit adjusts between trials to improve its ability to form preferred decision-spaces (advantage) while avoiding large changes to activity during decisions (cost). Advantage is represented mathematically as a weighted sum of probabilities that the possible decision-spaces form, given the circuit elements $\{X_1, X_2, \ldots, X_n\}$ take a specified set of activities $x_1, x_2, \ldots, x_n$ (eq. (19)). Cost is represented as the distance between the circuit activity used during decision-making and a "baseline" circuit activity that the circuit takes outside of decision-making (eq. (20)). Net advantage is the difference between advantage and weighted cost (eq. (21)). The circuit adjusts between trials in the direction of maximal change in net advantage (eq. (22)). See ***Movement of Circuit Activity Across Multiple Trials*, Methods**.

(legend continued on next page)

Extended Data Fig. 8 continuation

b,c, A possible explanation for the observation that in certain psychiatric disorders (post-traumatic stress disorder and substance use disorder) exposure to a traumatic event or drug can lead to increasingly altered choice after an extended period of nonexposure or abstinence. This can be modeled as adaptation in the circuit to process a rare decision-dimension that was necessary to process, for example, the traumatic event or drug. In the simulation, the circuit's preference for various decision-spaces is updated over time based on their success in making choices according to environmental stimuli. In the disorder resilient scenario (**b**), the circuit is first exposed to the traumatic event or drug at the dashed line. The "red" decision-space is not formed as often as are other decision-spaces (gray lines). In contrast, even short periods of exposure to the traumatic event or drug can impact future decision-spaces, resulting in vulnerability (**c**). Circuit activity adjusts until, when the traumatic event or drug reappears, the "red" decision-space forms frequently. This result may explain incubation of fear (in post-traumatic stress disorder) or craving (in a substance use disorder), where symptoms emerge only after a period of weeks after exposure. Differences in response post-incubation lead to modeled vulnerability or resilience.

d, Circuit activity morphs over time in response to decision-making needs.

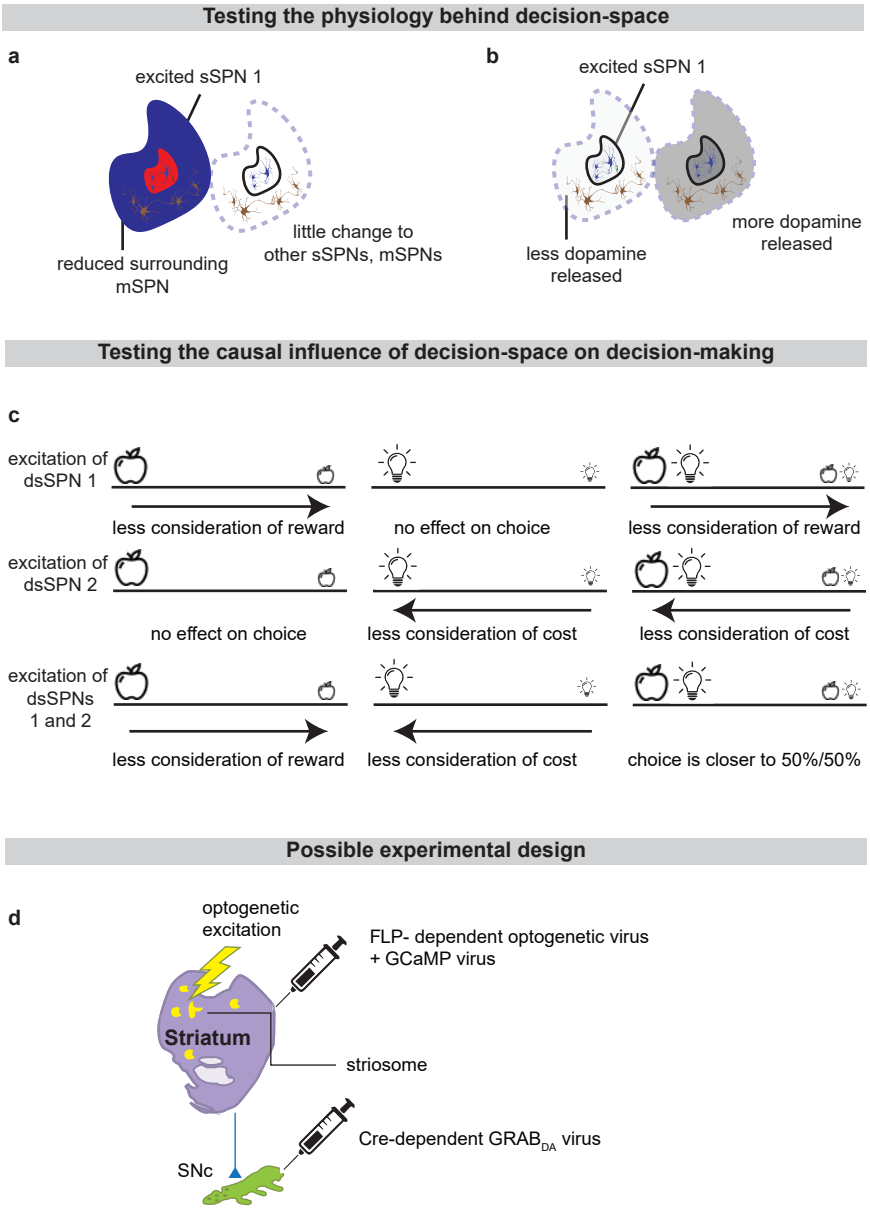
(19) $$\text{advantage}(X_1=x_1,\ X_2=x_2,\ldots,X_n=x_n) = \sum_{l=1}^{2^q} \text{score}_l \cdot \text{P}(\text{space}_l \mid (X_1=x_1,\ X_2=x_2,\ldots,X_n=x_n))$$

(20) $$\text{cost}(X_1=x_1,\ X_2=x_2,\ldots,X_n=x_n) = \left\| \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^\intercal - \begin{bmatrix} x_{1,\text{baseline}} & x_{2,\text{baseline}} & \cdots & x_{n,\text{baseline}} \end{bmatrix}^\intercal \right\|_2$$

(21) $$\text{net advantage}(X_1=x_1,\ldots,X_n=x_n) = \text{advantage}(X_1=x_1,\ldots,X_n=x_n) - \text{constant} \cdot \text{cost}(X_1=x_1,\ \ldots,X_n=x_n)$$

(22) $$\frac{\Delta \begin{bmatrix} x_{1,\text{baseline}} & x_{2,\text{baseline}} & \cdots & x_{n,\text{baseline}} \end{bmatrix}^\intercal}{\text{trial}} = \text{rate} \cdot \nabla \text{net advantage}(X_1=x_1,\ldots,X_n=x_n)$$

Extended Data Fig. 9



Extended Data Fig. 9: Proposed experiments that might add support to our model.

a, The decision-space model assumes that when an sSPN subpopulation is active, the corresponding decision-dimension is unlikely to be used to form the decision-space. Thus, activating an sSPN subpopulation (red) should lead to lower activity in neighboring mSPNs (blue) and there should be lower variance in their activities. This would support our hypotheses that sSPNs bias a nearby mSPN subpopulation towards being used or excluded during the formation of decision-space.

b, Meanwhile, we would expect to find less dopamine (light gray versus darker gray) released to the sSPN subpopulation and neighboring mSPNs than to other SPNs. This would support our hypothesis that the decision-space is formed via selective dopamine release.

(legend continued on next page)

Extended Data Fig. 9 continuation

c, We would expect sSPN stimulation to alter decision-making. An experiment could be designed asking rodents to choose rewards (apple) and/or costs (lights) in a T-maze (line, animal begins in the center). The rodents would be expected to selectively deprioritize informational dimensions corresponding to sSPN subpopulations that are stimulated. For instance, stimulating reward-responding sSPNs would be expected to lead to reduced consideration of reward when making the decision.

d, The hypotheses in **a-c** can be tested with tools available to neuroscientists. Similar to was performed by Lazaridis et al. (2024), DA-Cre-expressing mouse are crossed with striosomal-FLP-expressing mice. Three viruses are injected: 1) Cre-dependent $GRAB_{DA}$ into daSNC (for instance, similar to Sun et al. (2020))[81], 2) FLP-dependent optogenetic virus into the striatum, 3) general G-CAMP virus into the striatum. Striosome and matrix are then recorded using two-photon microscopy, similar to, for example, Bloem et al. (2022). Four colors are used to: 1) identify dopamine via the $GRAB_{DA}$ virus 2) identify all striosomal and matrix neurons via the GCaMP virus, 3) identify the contours of the striosomes via the FLP-dependent virus, and 4) stimulate the striosome. In a closed loop way, striosome neurons are identified that correspond to reward or cost, and the neurons that selectively respond to each are stimulated as rodents perform a T-maze task that was used by, for instance, Friedman et al. (2015).

SUPPLEMENTARY MATERIALS

Supplementary Note 1: Choice of circuit elements. Using a reductionist approach, we selected circuit elements *crucial to the regulation of dopamine by dorsomedial striosomes*. Thus, we prioritized the direct connection between striosomes and daSNC and the regulation of daSNC via GPe and GPi→LHb→RMTg.

Notably, we exclude from our model the substantia nigra reticula (SNr), another basal ganglia region that receives projections from the striosomes[24,95] and the subthalamic nucleus[96] and feeds back to daSNC neurons[97], forming an additional secondary striosome→daSNC connection. In the context of our model, sSPN→SNr→daSNC helps to determine *how many* decision-dimensions are included in the parallel direct and indirect pathway decision-spaces, in conjunction with the modeled operation assigned to the GPi→LHb→RMTg pathway. The connection through SNr, by introducing a striosome→SNr→daSNC→striosome loop, would add an additional set of dynamics on a longer time scale than the modeled striosome→daSNC→striosome loop. The analysis using ***Instance 1*** of the model (full connectivity, no dynamics), however, would remain similar, with similar selective effects of striosomes, GPi, LHb, RMTg, and daSNC on the decision-space and choice.

Other regions that might be included in a future expanded model are the central nucleus of the amygdala, the paraventricular thalamic nucleus, the rhomboid, and the paratenial thalamic nucleus, each of which provides input to the striosomes more so than matrix[70,98–101]. In the context of our model, these connections would help to influence *which* decision-dimensions are included in decision-space, in conjunction with cortical neurons, FSIs, sSPNs, and daSNC neurons.

More broadly, our model does not include certain brain regions that are relevant to decision-making such as the dorsolateral striatum, ventral striatum, dopaminergic neurons in the ventral tegmental area (VTA), subthalamic nucleus, and other basal ganglia regions. There are also brain regions outside the basal ganglia that are implicated in decision-making[102] that we do not consider.

We did not consider neuronal molecular heterogeneity or gene expression.

Supplementary Note 2: Choice of cortical inputs to the striatum. We focus on the processing of information by the striatum rather than the specific information encoded by each region of the cortex. This allows us to retain focus on the processing of information by the striatum. Additionally, our model assumes that cortical inputs are roughly similar between the striosomes and matrix. This assumption allows us to assign the same decision-dimensions to striosome subpopulations as to matrix subpopulations. As shown in **Table 2**, in reality, several striatum-projecting cortical regions have been found to project more strongly to either striosomes or matrix[2,67,69], although the experimental data on many of these regions is mixed[70]. However, we suggest that there are somewhat similar representations of decision-dimensions in sSPNs and mSPNs. Computationally, as shown in **Extended Data Fig. 4**, average population activities can successfully encode information mapped to the principal components of cortical activity even when cortex→SPN connectivity is sparse and different cortical neurons project to each SPN. So, provided that the cortical information encoded in striosome- and matrix-projecting regions is

somewhat similar, the decision-dimensions formed by the population should, on average, be similar. To speculate on the function role of the different connectivity, it could be that the regions that are more mSPNs-projecting relay more detailed representations of information, while the regions that are more sSPNs-projecting relay information about how many decision-dimensions to use (a similar effect to the b_{sSPN} term in eq. (1)). This hypothesis might align with functional differences in the regions that project mostly to striosomes versus matrix. For instance, the prefrontal cortex, associated with complex cognitive function, projects mostly to striosomes and inhibits striosomes during conflict[11], perhaps suggesting a functional role in increasing the dimensionality of the decision-space based on internal representations of task difficulty. Similarly, the orbitofrontal cortex, associated with high-level decision-making schema[103] and hyperactive during anxiety, projects mostly to striosomes and may communicate contextual information related to the decision.

Supplementary Note 3: Modulation of SPNs by dopamine during the decision to form decision-space. We make several hypotheses about dopamine in order to form our model. First, we hypothesize that dopamine plays a role in modulating SPN activity during decision-making. This follows based on the existing work linking dopamine to second-by-second modulation of SPN firing[104] including work that demonstrates an opposite effect in striosome versus matrix[20]. A recent experiment demonstrated that midbrain dopamine neurons selectively modulate a subpopulation of SPNs[89], perhaps corresponding to a decision-dimension (see **Table 7**). Second, we hypothesize that this modulatory effect occurs variably over time and is most pronounced when new important cues appear. This aligns with the literature on RPE[105]. Third, we hypothesize that dopamine modulation should only occur in complex tasks (that require one or more decision-dimension), compatible with evidence that suggests minimal effect of dopamine release on SPNs in simple tasks[89]. Fourth, we hypothesize that dopamine plays an important role in largescale sSPN modulation of mSPN compared to including collaterals, interneurons (choline acetyltransferase interneurons, somatostatin expressing interneurons, and parvalbumin-positive interneurons cross compartment boundaries but likely have primarily local effects[106]), or other possible connections through other brain regions.

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Alexander Friedman (afriedman@utep.edu)

Materials availability

Code used to construct, analyze, and test the model is deposited to https://github.com/dirkbeck/DM_space_model.

Code used to analyze neural data from the Corticostriosomal Circuit Stress Experiment database is deposited to https://github.com/dirkbeck/DM_space_model.

Data from the Corticostriosomal Circuit Stress Experiment data, prepared for use in the current paper, is deposited to https://doi.org/10.7910/DVN/SMKW0I.

Data and code availability

- This paper analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.
- All original code has been deposited at https://github.com/dirkbeck/DM_space_model and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHODS

Outline.

- ***Decision-dimensions and decision-space.*** Explanation of the foundational concept of the model.
- ***Analyzed instances of the model.*** We conduct our analysis using three instances of the conceptual model. In the following sections, we formally define the model in each instance, then detail the methods behind our related analyses.
    - ***Instance 1: full connectivity and feedforward.*** Related to **Figs. 1-3,5, Extended Data Figs. 1-3,7**. Used to link neural activity, the decision-space, and choice.
    - ***Modeled Circuit Manipulation using Instance 1.***
    - ***Instance 2: sparse connectivity and feedforward.*** Related to **Extended Data Fig. 4**. Used to demonstrate how a large network might encode the decision-space.
    - ***Modeling SPN encoding of data, using Instance 2.***
    - ***Instance 3: full connectivity and dynamics.*** Related to **Fig. 4, Extended Data Fig. 6**. Used to demonstrate how the decision-space might form over time.
    - ***Modeling time-variant input, using Instance 3.***
- ***Movement of Circuit Activity Across Multiple Trials.*** An extension of the model to view possible changes of the circuit between trials in the context of decision-space. Related to **Fig. 6, Extended Data Fig. 8**.
- ***Rationale for the computational framework.*** Reasoning behind our modeling strategies.
- ***Inferring decision-space from SPN activity and choice.*** A method we designed in which decision-space can be inferred from experimental data. Related to **Extended Data Figs. 1q,r**.
- ***Testing the Model Through Analysis of Neural Data.*** Analysis of neural data which supports our model. Related to **Figs. 2e,f, Extended Data Figs. 2a-n, 3a-p.**

Decision-dimensions and decision-space.

The physiologies of the circuit elements produce two abstractions which we use, for convenience, throughout or work:

- A *decision-dimension* is an axis of the coordinate system with which SPNs (dsSPNs, isSPNs, dmSPNs, imSPNs; see **Table 1** for anatomical definitions) encode data projected from the cortex. A decision-dimension is equivalent to a principal component of cortical activity. In our analysis, separate groups of SPNs encode data along the first, second, third, and fourth principal components. (Arbitrarily, we do not consider principal components beyond the first four). Each of the dsSPN/isSPN/dmSPN/imSPN subgroups have neurons corresponding to each of the four principal components.
- The *decision-space* is the mathematical space formed by mSPNs (both dmSPNs and imSPNs) after dopamine signaling from daSNC. Modeled dopamine signaling determines whether to include or exclude neurons encoding each decision-dimension during a decision. Thus, from the mathematical space formed by all decision-dimensions, a

mathematical subspace (i.e. the “decision-space”) is formed with which to define action values.

We use the prefix “decision” because the circuit uses decision-space, formed from a basis of decision-dimensions, to define action values during decision-making.

Analyzed Instances of the Model.

The general case of the model (although not formally used for analysis) is a dynamic network of cortical neurons, FSIs, dsSPNs, isSPNs, GPi, LHb, daSNCs, dmSPNs, imSPNs, and mSPN-projecting neurons which encode action values.

We conduct our analyses using three instances of this general case, which are each equivalent to the general case under the specific conditions we outline. The three instances, tailored to our various analyses, each allow for a different mathematical simplification. This allows us to conceptually and formally define the instances individually in a way that is intuitive and relates directly to our analyses.

1) Instance 1 has full cortex→FSI→SPN connectivity and constant activity in each circuit element throughout the decision. In this instance, the model can be defined equivalently using a smaller set of network elements and a feedforward network. See ***Instance 1: full connectivity and feedforward.***
2) Instance 2 has constant activity in each circuit element throughout the decision. In this instance, the model can be defined equivalently using a feedforward network. See ***Instance 2: sparse connectivity and feedforward.***
3) Instance 3 has full cortex→FSI→SPN connectivity. In this instance, the model can be defined equivalently using a smaller set of network elements. See ***Instance 3: full connectivity and dynamics.***

Instance 1: full connectivity and feedforward.

In this section, we describe the instance of the model where each cortical neuron projects to each FSI, each FSI projects to each SPN (for dsSPN, isSPN, dmSPN, imSPN), and each cortical neuron projects to each SPN. Additionally, cortical input to the system does not change over time, and the activities of other circuit elements do not decay over time.

This instance of the model leads to a convenient formation of the model as a circuit of fewer elements (one FSI, one dsSPN, isSPN, dmSPN, and imSPN per the four decision-dimensions), and no time component. In this section, we frame this instance mathematically and then describe our related analysis.

Input: cortical activity.

During a decision, a vector of cortical input $\mathbf{x}_P \in \mathbb{R}^{p \times 1}$ enters each pathway P in the network ($\mathbf{x}_{\text{direct}}$ to direct pathway SPNs and $\mathbf{x}_{\text{indirect}}$ to indirect pathway SPNs). The elements of $\mathbf{x}_P$ are the activities of p cortical neurons. Each neuron encodes a different sensory input.

Outputs.

We use this instance of the model to examine: 1) the activities of the circuit elements depending on the activities of other circuit elements (**Fig. 1, Extended Data Figs. 1a-i**); 2) the modulation of mSPNs by dopamine (i.e. decision-space, **Figs. 1a,b**); 3) action values given decision-space (**Extended Data Figs. 1j-l**), and 4) choice given action values (**Extended Data Figs. 1m-p**).

1) The activities of circuit elements during a decision are related to each other based on anatomically realistic connections (eqs. (1),(2),(3),(5),(23)).
2) A decision-space is formed probabilistically. The probability a given decision-dimension i being used during a decision is equivalent to the activity of daSNC (see eq. (2)), which ranges from 0 to 1. Probabilities are realized in the connection from daSNC to mSPN (see eq. (3)), when each decision-dimension is probabilistically assigned a weight (in most analyses, either 0 or 1). Decision-space is defined as the space formed from the basis of decision-dimensions that were not assigned a weight of 0.
3) Action value is derived based on mSPN activity during a decision.
4) Choice is derived from action values. Action values are treated as Merton processes[107] using eq. (6). Several possible actions are assigned action values and the corresponding process that hits the threshold first is enacted.

Defining FSI activity.

FSI activity, c_P, is set to the magnitude of $\mathbf{x}_P$ for each pathway, multiplied by a weight of cortex→FSI connection a_{FSI}, plus an additive shift b_{FSI}:

(23) $$c_P = a_{FSI} \cdot \left\| \mathbf{x}_P \right\| + b_{FSI}$$

where:

- c_P is relative activities of FSIs that project to SPNs of pathway P (activity arb. u.)
- a_{FSI} is the weight of cortex→FSI connection. Similar for both P. (dimensionless)
- $\mathbf{x}_P$ is the activities of cortical neurons that project to SPNs of pathway P. (activity arb. u.)
- b_{FSI} affects the relative activity of all sSPN neurons. Similar for both P. (activity arb. u.)

In the current instance of the model, there are 2 FSIs, one that receives input from $\mathbf{x}_{\text{direct}}$ and projects to dsSPNs and dmSPNs, and the other than receives input from $\mathbf{x}_{\text{indirect}}$ and projects to isSPNs and imSPNs.

For use in our analysis, see https://github.com/dirkbeck/DM_space_model/blob/main/algorithmic_model.m.

Defining sSPN activity.

To get the activities of sSPNs in each pathway, $\mathbf{x}_P$ is normalized via division by c_P and multiplied by $\mathbf{W}_P \in \mathbb{R}^{p \times q}$, which linearly transforms and reduces cortical input from the p-dimensional coordinate space of cortex to the smaller q-dimensional coordinate space of sSPN. In the sSPN coordinate space, each coordinate is a principal component of a training set of historical cortical input across n time steps $\mathbf{X}_P \in \mathbb{R}^{n \times p}$ (uncorrelated, for simplicity). For each pathway, $\mathbf{W}_P$ contains the truncated first q columns (corresponding to the first q principal components) of $\mathbf{W}_{\text{full},P} \in \mathbb{R}^{p \times p}$ after the decomposition $\mathbf{X}_P \mathbf{X}_P{}^{\intercal} = \mathbf{W}_{\text{full},P} \mathbf{\Lambda} \mathbf{W}_{\text{full},P}{}^{\intercal}$ is made to obtain the full principal component matrix. Experimental work has revealed dimensionality reduction the order of ~100 times from cortex to SPNs[108], so $q \ll p$. Note that in our analysis using the current instance of our model, $\mathbf{X}_P$ is not explicitly generated because we specify the inputs to the system in terms of the coordinate space of decision-dimensions.

For each pathway, the components of an sSPN activity vector $\mathbf{s}_{\mathbf{sSPN},P} \in \mathbb{R}^{q \times 1}$ each correspond to the activity of an sSPN circuit element. A constant b_{sSPN}, used in analyses where modeled sSPN activity is stimulated or inhibited, adjusts overall sSPN activity:

(1) $$\mathbf{s}_{\mathbf{sSPN},P} = \frac{1}{c_P} \mathbf{W}_P{}^{\mathbf{T}} \mathbf{x}_P + b_{\text{sSPN}}$$ (copied from **Results** for convenience)

where:

- c_P is the relative activity of the FSI projecting to SPNs of pathway P (activity arb. u.)
- $\mathbf{W}_P$ is a matrix of weights from cortical neurons to SPNs of pathway P. Each column is equivalent to a principal component of cortical activity. (dimensionless)
- $\mathbf{x}_P$ is the activities of cortical neurons that project to SPNs of pathway P. (activity arb. u.)
- b_{sSPN} affects the relative activity of all sSPN neurons (activity arb. u.)

In the current instance of the model, activities are defined based on a feedforward network, so the simplification is made that sSPN activities are not affected by daSNC activities.

In the current instance, there is one sSPN per decision-dimension per pathway. So, there are q dsSPNs and q isSPNs. The dsSPNs receive input from $\mathbf{x}_{\text{direct}}$ and c_{direct}. The isSPNs receive input from $\mathbf{x}_{\text{indirect}}$ and c_{indirect}.

For use in our analysis, see https://github.com/dirkbeck/DM_space_model/blob/main/algorithmic_model.m.

Defining GPi, LHb, and RMTg activities.

The GPi→LHb→RMTg→daSNC pathway performs a series of operations which influence RMTg activity RMTg, which is an input to daSNC activity in eq. (2). Weights $\mathbf{w}_{\text{GPi},P} \in \mathbb{R}^{2q\times 1}$, not necessarily positive, are combined with the activities of the q dsSPNs and q isSPNs, forming scalar representations of dsSPN or isSPN activity. z_{GPi}, z_{LHb}, and z_{RMTg} terms reflecting the activities of those circuit elements are combined with this scalar representation:

$$\mathrm{RMTg} = z_{\text{RMTg}} + z_{\text{LHb}} + z_{\text{GPi}} \cdot \mathbf{w}_{\text{GPi}} \cdot \begin{bmatrix} \mathbf{s}_{\text{sSPN, direct}} \\ \mathbf{s}_{\text{sSPN, indirect}} \end{bmatrix} \quad (5)$$

(copied from **Extended Data Fig. 1** for convenience)

where:

- z_{RMTg} is an additive shift that affects relative RMTg activity (activity arb. u.)
- z_{LHb} is an additive shift that affects relative LHb activity (activity arb. u.)
- z_{GPi} is a coefficient that affects relative GPi activity (activity arb. u.)
- $\mathbf{w}_{\text{GPi}}$ is the weights of connection from sSPNs of pathway P to the GPi neuron (dimensionless)
- $\mathbf{s}_{\text{sSPN},P}$ is the activity of sSPNs corresponding to decision-dimension i and pathway P (activity arb. u.)

This pathway contains one GPi element, one LHb element, and one RMTg element, as visualized in **Fig. 1a**. All sSPN elements project to GPi. GPi activity is an input to LHb, which after the z_{LHb} addition, is an input to RMTg, which itself has a z_{RMTg} addition. For simplicity, these series of operations are presented together in eq. (5).

For use in our analysis, see https://github.com/dirkbeck/DM_space_model/blob/main/algorithmic_model.m.

Defining daSNC activity.

daSNC neurons incorporate the output of the GPi→LHb→RMTg→daSNC pathway with direct inputs from sSPN elements. There are q sSPN elements of each pathway and q daSNC elements corresponding to each pathway. For each pathway, the ith sSPN element connects to the ith daSNC element, but not to other daSNC elements (see **Fig. 1a**). These connections have weights $w_{\mathrm{sSPN}\to\mathrm{daSNC},i,P}$ for $i=1,2,\ldots,q$. RMTg, on the other hand, connects to each daSNC element. The output of the ith daSNC element, constrained to between 0 and 1 via a logistic function, captures the importance of a single decision-dimension:

(2) $$\mathrm{daSNC}_{i,P}=\frac{1}{1+\exp\left(w_{\mathrm{sSPN}\to\mathrm{daSNC},i,P}\cdot \mathrm{s}_{\mathbf{sSPN},i,P}+\mathrm{RMTg}-z_{\mathrm{daSNC},i,P}\right)}$$

(copied from **Results** for convenience)

where:

- $w_{\mathrm{sSPN}\to\mathrm{daSNC},i,P}$ is the weight of connection from the sSPN corresponding to decision-dimension i and pathway P to the daSNC corresponding to decision-dimension i and pathway P . The weight is fixed in this instance of the model. (dimensionless)
- $\mathrm{s}_{\mathrm{sSPN},i,P}$ is the activity of sSPNs corresponding to decision-dimension i and pathway P (activity arb. u.)
- $z_{\mathrm{daSNC},i,P}$ is an additive shift applied to the daSNC neuron corresponding to decision-dimension i and pathway P (activity arb. u.)
- RMTg is RMTg activity, as defined in eq. (5). (activity arb. u.)

This pathway is modeled using one neuron per decision-dimension per pathway. So, there are q daSNC neurons that each receive projection from a dsSPN, and q daSNC neurons that each receive projection from an isSPN. daSNC elements also receive input from RMTg.

For use in our analysis, see https://github.com/dirkbeck/DM_space_model/blob/main/algorithmic_model.m.

Defining mSPN activity and decision-space.

In each pathway, decision-space is formed probabilistically. The conversion from daSNC activity to realization of decision-space occurs in the connections from daSNC to mSPN. There are q daSNC elements corresponding to each pathway and q mSPN elements, and, in each pathway, the ith daSNC element connects to the ith mSPN element, but not to other mSPN elements (see **Fig. 1a**).

Like sSPNs, mSPNs encodes the cortical input normalized by an FSI and is transformed to a coordinate space of the first q principal components. The difference is that for each of dmSPNs and imSPNs, a diagonal matrix $\mathbf{S}_P \in \mathbb{R}^{q \times q}$ is multiplied by the cortical input after transformation:

(3) $$\mathbf{s}_{\mathrm{mSPN},P} = \frac{1}{c_P}\,\mathbf{S}_P \mathbf{W}_P{}^{\mathrm{T}} \mathbf{x}_P$$ (copied from **Results** for convenience)

where:

- c_P is the relative activity of the FSI projecting to SPNs of pathway P (activity arb. u.)
- $\mathbf{S}_P$ is a diagonal matrix that applies dopamine release (via daSNC activity) to mSPN activity in pathway P. (dimensionless)
- $\mathbf{W}_P$ is a matrix of weights from cortical neurons to SPNs of pathway P. Each column is equivalent to a principal component of cortical activity. (dimensionless)
- $\mathbf{x}_P$ is the activities of cortical neurons that project to SPNs of pathway P. (activity arb. u.)

In the current instance, there is one mSPN per decision-dimension per pathway. So, there are q dmSPNs and q imSPNs. The dmSPNs receive input from $\mathbf{x}_{\mathrm{direct}}$ and c_{direct}. The imSPNs receive input from $\mathbf{x}_{\mathrm{indirect}}$ and c_{indirect}.

The diagonal elements of $\mathbf{S}_P$ are set probabilistically to either 1 (dimension in decision-space) or 0 (dimension not in decision-space) such that $\mathrm{P}\left(\mathbf{S}_{\mathbf{P},\mathrm{ii}} = 1\right) = \mathrm{daSNC}_{i,\mathrm{P}}$.

Thus, in the portions of our analysis where we set the activities of the q daSNC elements to be equal, the decision-dimensions each have the same probability of being included in decision-space, i.e. $\mathrm{daSNC}_1 = \mathrm{daSNC}_2 = \ldots = \mathrm{daSPN}_q = d$. In this case, we treat the probability of a certain decision-space dimensionality forming as a binomial distribution:

(18) $$\mathrm{P}(m \text{ DM-dimensions used to form DM-space}) = \binom{q}{m} d^m (1-d)^{q-m} \quad \text{for } m = 0, 1, \ldots, q$$

(copied from **Extended Data Fig. 8** for convenience)

where:

- q is the number of possible decision-dimensions
- d is the (equal) probability that each decision-dimension is used to form decision-space

Defining action value.

Action value (or, in the indirect pathway, inaction value) $v_{j,P}$ for each of k potential actions is defined based on the activities of dmSPNs (or imSPNs). During this process, elements of a coefficient matrix $\boldsymbol{\beta}_P \in \mathbb{R}^{k \times q}$ are applied to mSPN activities for each decision-dimension, action, and pathway. Bias $\alpha_{j,P}$ is subtracted. Below, $\boldsymbol{\beta}_{j,P}$ is used to indicate row j of $\boldsymbol{\beta}_P$.

(4) $$v_{j,P} = \frac{1}{1 + \exp\left(-\boldsymbol{\beta}_{j,P}\, \mathrm{s}_{\mathrm{mSPN},P} - \alpha_{j,P} \right)}$$ (copied from **Results** for convenience)

where:

- $\boldsymbol{\beta}_{j,P}$ is a matrix of weights from dmSPNs to downstream action value encoding neurons for the direct pathway, or imSPNs to downstream inaction value encoding neurons for the indirect pathway. (dimensionless)
- $\mathrm{s}_{\mathrm{mSPN},P}$ is the activity of sSPNs corresponding to decision-dimension i and pathway P (activity arb. u.)
- $\alpha_{j,P}$ is an additive shift corresponding to the neuron encoding action j for the direct pathway or inaction j for the indirect pathway. (activity arb. u.)

There is one neuron encoding each $v_{j,P}$. So, there are k neurons encoding action values and k neurons encoding inaction values. Each of these neurons receives projection from all mSPNs of the corresponding pathway.

Defining choice.

k Merton process[107] are run to determine whether each action should be taken, and another k to determine whether each action should be refrained from. Progress to choice for each action (or inaction), $Y_{j,P}$, is related to its corresponding action (or inaction) value $v_{j,P}$ and an uncorrelated Brownian component $\mathrm{d}W_{j,P}$ scaled by a coefficient σ.

(6) $$\mathrm{d}Y_{j,P} = v_{j,P}\,\mathrm{d}t + \sigma\,\mathrm{d}W_{j,P},\ Y_{j,P}(t=0) = 0,\ \text{where } W_{j,P} \text{ is a standard Wiener process}$$

(copied from **Extended Data Fig. 1** for convenience)

where:

- $Y_{j,P}$ is the progress to enaction of action j in the direct pathway, and progress to refraining from action j in the indirect pathway.

- $v_{j,P}$ is the action (or inaction) value corresponding to action j and pathway P. (activity arb. u.)
- σ is the coefficient of noise.

The time it would take to enact action j, $t_{\text{action},j}$, is defined as the first hit time of a threshold h for process j of the direct pathway:

$$t_{\text{action},j} = \min_t \left\{ t \mid Y_{j,\text{direct}} \geq h \right\} \quad (7)$$

(copied from **Extended Data Fig. 1** for convenience)

The time it takes to exclude action j from consideration, $t_{\text{inaction},j}$, is calculated similarly using the indirect pathway:

$$t_{\text{inaction},j} = \min_t \left\{ t \mid Y_{j,\text{indirect}} \geq h \right\} \quad (8)$$

(copied from **Extended Data Fig. 1** for convenience)

The enacted action is the first to reach h, given that the corresponding inaction process has not first reached h:

$$\text{action} = \arg\min_{j \in J} \left(Y_j(t_{\text{action},j}) \right), \text{ where } J \text{ is the subset of actions s.t. } t_{\text{action},j} < t_{\text{inaction},j} \quad (9)$$

(copied from **Extended Data Fig. 1** for convenience)

where:

- $Y_{j,P}$ is the progress to enaction of action j in the direct pathway, and progress to refraining from action j in the indirect pathway.
- t is time (s)
- h is a threshold at which an action is considered taken (progress to decision arb. u.)

In our analysis, we run simulations using a constant time step discretization of eq. (6).

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/weiner_process_model.m.

Modeled Circuit Manipulation using Instance 1.

To get a sense of the functional role of the circuit elements, we conducted sensitivity analyses by changing parameters in the model individually and determining their effect on the activities of other circuit elements, decision-space formation, action values, and/or choice.

Common parameters.

The values specified here, arbitrarily chosen, are used in the analyses in ***Instance 1*** unless otherwise indicated:

- throughout, $k = 4$
- in eq. (23): $a_{\mathrm{FSI}} = 1$
- in eq. (23): $b_{\mathrm{FSI}} = 0.5$
- in eq. (1): $b_{\mathrm{sSPN}} = 0$
- in the inputs to eq., $q = 4$
- in eq. (5): $z_{\mathrm{GPi}} = 1$
- in eq. (5): $z_{\mathrm{LHb}} = 0.5$
- in eq. (5): $z_{\mathrm{RMTg}} = 0.5$
- in eq. (2): $w_{\mathrm{sSPN} \to \mathrm{daSNC}, i, P} = 1$ for all i and P
- in eq. (2): $z_{\mathrm{daSNC}, i, P}$ for all i and P
- in eq. (4): $\boldsymbol{\beta}_{\mathrm{direct}} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$,

  whose rows correspond to, for example: turning left, turning right**,** turning around**,** wandering; and whose columns correspond to, for example: a reward-predominant decision-dimension 1, a cost-predominant decision-dimension 2, a novelty-predominant decision-dimension 3, and a location-predominant decision-dimension 4. The coefficients model a T-maze where a choice is made to turn right or left based on relative values of cost and reward.

- in eq. (4): $\alpha_{j,P} = -3$ for all j and P
- in eq. (6): $\sigma = 1$
- in eq. (7), (8): $h = 2$

Effect of reward/costs on LHb/RMTg/daSNC activity.

In **Extended Data Figs. 1e,f,** we modeled the effect of incrementing reward or cost on the activities of LHb, RMTg, and daSNC.

The inputs enter the model circuit in two ways: 1) reward and cost are mapped to decision-dimensions; and 2) cost level leads to changes in LHb and RMTg activities, similar to what has been demonstrated in experimental work[75,82,109–111]. The modeled LHb and RMTg responses to cost are proportional to cost level with an arbitrary coefficient (set to 1 for LHb and 0.9 for RMTg for the purposes of plotting).

The modeled results show that the mean activity of a daSNC subpopulation encoding reward-predominant data responds positively to increases in reward and negatively to decreases in reward, similar to experimental evidence[112]. LHb and RMTg respond negative linearly to reward level and positive linearly to cost level, similar to experimental evidence[109,113]. Sudden changes in reward or cost level, therefore, lead to shifts in activities that track changes to expectations of future reward or cost value, including reward or cost currently received, i.e. reward or cost prediction error.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/model_overview/GPi_LHb_RMTg_DA_model.m.

Effect of LHb/RMTg/daSNC activity on decision-space.

In **Extended Data Figs. 1g-i**, we modeled the effect of incrementing GPi, LHb, RMTg, or daSNC activity on the type of decision-space formed during a decision.

In the plotted analysis, we altered z_{GPi} in eq. (5), z_{LHb} in eq. (5), z_{RMTg} in eq. (5), and $z_{\mathrm{daSNC},i,P}$ (uniform change for all i, a single pathway is considered) in eq. (5) such that they took 10 values incremented from 0 to 1. Parameters not altered took default values (see **Common parameters**).

We also examined the role of each component in decision-space formation through the perspective of a series of steps, each carried out by a different circuit element. For this analysis, we substituted eq. (5) into eq. (2) and altered each parameter in turn. The plots illustrate the value of $\mathrm{daSNC}_{i,P}$ if the other parameters were set to 1 (z_{GPi}) or 0 (z_{RMTg}, $z_{\mathrm{daSNC},i,P}$). b_{LHb} is set to 0.5 (control), -5 (lesioned LHb), or 5 (stimulated LHb).

See **Tables 6,7** for alignment to the experimental literature.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/model_overview/GPi_LHb_RMTg_DA_model.m.

Effect of sSPN activity on decision-space.

In the analysis plotted in **Fig. 2a,** we incremented b_{sSPN} in eq. (1) and, for each increment, recorded daSNC_i in eq. (2). Then, using the approach in eq. (6), we converted the probability that one decision-dimension is used in the formation of decision-space to the probability that a decision-spaces of a certain dimensionality is formed.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/model_tests/friedman2015optogeneticmanipulation.m.

Effect of decision-space on choice.

In the analysis plotted in **Fig. 2b**, we changed which decision-space was formed by mSPNs and measured choice.

The excitation group was modeled using a non-dimensional decision-space (dopamine→mSPN weights of 0 reward-predominant decision-dimension, 0 cost-predominant dimension). The control group was modeled using a 1D direct pathway decision-space (dopamine→mSPN weights of 0.5 reward-predominant dimension, 0 cost-predominant dimension). The inhibition group was modeled using a 2D direct pathway decision-space (dopamine→mSPN weights of 1 reward-predominant dimension, 1 cost-predominant dimension). The modeled T-maze task was a choice between reward=2, cost=1 (high reward, high cost) and reward=1, cost = 0.5 (low reward, low cost). 20 simulations were run per modeled subject for 100 subjects. Other parameters for forming decision-space and calculating action value are set to their defaults (see **Common parameters**). For simplicity, the indirect pathway is not modeled in this analysis.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/model_tests/friedman2015optogeneticmanipulation.m.

In the analysis in **Figs. 3b,c**, we modeled changes to decision-space and choice after stress.

Here, modeled control rodents made decisions using a 2D direct pathway decision-space formed from reward-predominant and cost-predominant decision-dimensions. This correspond mathematically to a truncation of β_{direct} (see eq. (4) and **Common parameters**) to two columns. The first subset of modeled stress-group rodents made decisions without forming direct pathway decision-space. This corresponds to an elimination of β_{direct} such that action value is defined purely based on priors ($\alpha_{j,\mathrm{direct}}$ in eq. (4)). The second subset made decisions without forming direct pathway decision-space until they reached a critical threshold, beyond which they formed a 1D direct pathway decision-space with a reward-predominant dimension. Action values are derived for the three groups across multiple reward and cost combinations (**Fig. 3c**) via eqs. (3) and (4). Then choices are modeled using eqs. (6), (7), and (9) across 2000 simulations per group for each reward concentration (each incremented from 0 to 1 arbitrary units, 7 increments). Cost concentration is set to 0.5 arbitrary units (set at this level to resemble the steepness of increase in the experimental psychometric function). Default parameters are used for action value

formation and the Merton process model. For simplicity, the indirect pathway is not modeled in this analysis. **Fig. 3b** plots the averages of the simulations.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/disorder_hypotheses/Friedman2017_lowD_space.m.

In the analysis plotted in **Figs. 3d,e**, we modeled the effect on choice of shifts in decision-space after a small cost is added to a reward (experimental data is plotted in **Extended Data Fig. 3j**).

Rodents in the only-reward task were modeled as forming a lower-dimensional direct pathway decision-space (decision-dimension 1 weight = 0.5, decision-dimension 2 weight = 0.2) while animals in the reward-and-cost task formed a higher-dimensional direct pathway decision-space (decision-dimension 1 weight = 1, decision-dimension 2 weight = 0.5).

To do this, we truncated β_{direct} (see eq. (4) and **Common parameters**) to two columns or derived action value purely based on priors ($\alpha_{j,\text{direct}}$ in eq. (4)). A cortical input of reward = 0.7, cost = 0.3 is shown in the plots. For simplicity, the indirect pathway is not modeled in this analysis.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/disorder_hypotheses/altered_choice_after_space_transition.m.

In the analysis plotted in **Fig. 3h**, we modeled changes to choice after aging in young and old groups.

Here, we truncated β_{direct} (see eq. (4) and **Common parameters**) to two decision-dimensions, the first corresponding to a reward-predominant decision-dimension and the second corresponding to a cost-predominant decision-dimension. In the current analysis, the first row of β_{direct} corresponded to licking while the second row corresponds to performing a different action, e.g. movement. The licking action was assigned a larger prior, $\alpha_{1,\text{direct}}=0$, $\alpha_{2,\text{direct}}=-3$, due to the strong association developed in the rodents between the experimental apparatus and licking. For the modeled "learned, young" group, no decision-space is formed during the reward-cue task and a decision-space using only a cost-predominant decision-dimension is formed during the cost-cue task (i.e. $\mathbf{s}_{\text{mSPN,reward,direct}}=\begin{bmatrix}0\\0\end{bmatrix}$, $\mathbf{s}_{\text{mSPN,cost,direct}}=\begin{bmatrix}0\\1\end{bmatrix}$ in eq. (3)). For the modeled "learned, old" group, no decision-space is formed during the reward-cue task and a decision-space involving a cost-predominant decision-dimension is partially formed during the cost task ($\mathbf{s}_{\text{mSPN,reward,direct}}=\begin{bmatrix}0\\0\end{bmatrix}$, $\mathbf{s}_{\text{mSPN,cost,direct}}=\begin{bmatrix}0\\0.5\end{bmatrix}$). For the "not learned" group, a decision-space involving a cost-predominant decision-dimension is partially formed during both tasks ($\mathbf{s}_{\text{mSPN,reward,direct}}=\begin{bmatrix}0\\0.5\end{bmatrix}$, $\mathbf{s}_{\text{mSPN,cost,direct}}=\begin{bmatrix}0\\0.5\end{bmatrix}$). For simplicity, the indirect pathway is not modeled in this analysis.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/disorder_hypotheses/Friedman2020_lowD_space.m.

Effect of decision-space on sSPN-mSPN correlation

In the analysis plotted in **Fig. 2d,** sSPN-mSPN correlation is compared across decision-spaces with different dimensionality.

It is assumed in the plotted examples that a 1D decision-space is only formed from the first decision-dimension, a 2D decision-space is only formed from the first and the second, and a 3D decision-space is only formed from the first, second, and third. The analysis assumes a comparison of SPNs of the same pathway (that is, either dsSPN-dmSPN or isSPN-imSPN). For this analysis, eigenvalues of cortical activity are set to 2, 1, 0.5, 0.2, and 0.1, respectively. Weighted averages of example signals (left panel) and correlation for different decision-spaces (right panel) are formed using the identity that eigenvalues of principal components are equivalent to their variances.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/model_tests/ctx_sSPN_mSPN_coordinated_activity.m.

Effect of FSI activity on decision-space.

In the analysis plotted in **Fig. 3f**, we incremented FSI activity a_{FSI} in eq. (23) and determined the response of daSNC_i in eq. (2) (a single pathway is considered). The activity parameters related to other circuit elements were held constant (see **Common parameters**).

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/disorder_hypotheses/space_dimensionality_vs_FSI.m.

Effect of cortical SNR on choice.

In the analysis plotted in **Extended Data Figs. 5a-e**, cortical signal to noise ratio (SNR) is altered and the effect on choice is simulated.

Merton process simulations (see **Defining Choice**) are run across ten increments of reward and cost from -1 to 1 arbitrary units for a modeled cost-benefit conflict task. Parameters related to action value are set to their defaults and the T-maze task is used (see **Common parameters**). Here, “turn right” corresponds to receiving the reward and cost combination, while other actions correspond to receiving no reward and no cost. For simplicity, only the direct pathway is used to

influence choice. 100 simulations are run for each of 100 reward and cost combinations, and for each combination, choice is averaged.

The above process is replicated with changes to two sets of parameters. First, the effect of changes to decision-space were considered. A different $\mathbf{S}$ in eq. (3) was used depending on specified decision-space: $\mathbf{S}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$ for 1D decision-spaces, and $\mathbf{S}=\begin{bmatrix} & & 0 \\ & \mathbf{I}_2 & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$ for 2D decision-spaces. Second, changes to cortical noise were considered by adding i.i.d. Gaussian noise to $\mathbf{x}_{\text{direct}}$ (with mean 0 and standard deviation σ) at every time step, then recalculating action values in eq. (4) based on the mSPN activities calculated at that time step. Simulations were run for $\sigma = 1, 2, \ldots, 10$. Default parameters were used for calculating action value (see **Common parameters**).

Examples of single simulations at each level of reward and cost are shown for the $\sigma = 1$ (high cortical SNR) and $\sigma = 5$ (low cortical SNR) cases in **Extended Data Figs. 5a-d**. In **Extended Data Fig. 5e,** expected value is averaged across the 100 simulations for each noise level. Expected value here is defined as reward minus 0.75*cost (to add preference for reward compared to cost, coefficient is arbitrary) achieved across reward and cost levels. In the plot, SNR is set to the inverse of σ.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/dynamic_model_and_neural_net/cortical_snr.m.

Effect of dopamine on action/inaction values.

In the analyses plotted in **Extended Data Figs. 6c-f**, we measured the effect of high versus low dopamine on action and inaction values across a range of cortical inputs to the system.

We modeled a cost-benefit conflict task with increasing reward (scale of 0 to 1 arbitrary units, 100 increments) and constant cost (set to 0.25 arbitrary units). Experimental work has shown that dopamine increases direct pathway activity while decreasing indirect pathway activity and vice versa[42]. Therefore, we set coefficients relating to overall activity of the pathways oppositely: in the low dopamine case, the direct pathway coefficient was 0.1 arbitrary unit and the indirect pathway coefficient 5 arbitrary units; and in the high dopamine case, the indirect pathway coefficient was 5 arbitrary units and the direct pathway coefficient 0.1 arbitrary unit. These coefficients were multiplied by $\boldsymbol{\beta}_{\text{direct}}$ or $\boldsymbol{\beta}_{\text{indirect}}$ in eq. (4), increasing or decreasing the overall sensitivity of action value on data along cortical principal components. In the model, changes to dopamine also involved a change in decision-space: due to their opposite effects on mSPN activity, dopamine biases the direct pathway towards forming higher-dimensional decision-spaces and the indirect pathway towards forming lower-dimensional decision-spaces. For the purpose of this analysis, eq. (3) is reframed to incorporate the effects of dopamine in scaling action value score (A, set to

either 5 or 0.1 arbitrary units in our analysis) and changing decision-space (B, set to 1 arbitrary unit when dopamine is high and 0 when dopamine is low). Here, individual elements are referenced through subscripts based on their j th row and column corresponding to the reward or cost dimension.

(24) $$v_{j,\text{direct}} = \frac{1}{1+\exp\left(-A\cdot\left(\beta_{j,\text{reward,direct}} + B\cdot\beta_{j,\text{cost,direct}}\right) - \alpha_{j,\text{direct}}\right)}$$

(25) $$v_{j,\text{indirect}} = \frac{1}{1+\exp\left(-A\cdot\left((1-B)\cdot\beta_{j,\text{reward,indirect}} + \beta_{j,\text{cost,indirect}}\right) - \alpha_{j,\text{indirect}}\right)}$$

where:

- A is the multiplicative effect of dopamine released to mSPNs (dimensionless)
- B is the effect of dopamine on decision-space (dimensionless)
- β_j is the connection weight from mSPN to an action value neuron. Each $\beta_{j,\text{reward}}$ or $\beta_{j,\text{cost}}$ corresponds to an element of the connection weight matrix β_P.
- prior_j is an additive shift corresponding to the neuron encoding action j (activity arb. u.)

The plot in **Extended Data Fig. 6c** compares the high dopamine and low dopamine cases. The plot in **Extended Data Fig. 6d** shows a similar analysis but for changes in parameters: cost is fixed at 0.5 arbitrary units, and A is set to either 2 arbitrary units (corresponding to the pathway not disconnected) or 0 (corresponding to the pathway disconnected). The plots in **Extended Data Figs. 6e,f** show progress to action in the case where reward = 1 arbitrary unit and cost = 0.25 arbitrary unit for low versus high dopamine. Deliberation time distributions are formed by aggregating the deliberation times across the 100 simulations. For parameters used for subjective valuation and deliberation time simulation, see **Common parameters**.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/dynamic_model_and_neural_net/direct_vs_indirect_pathway_SV.m.

Effect of decision-dimensions on choice.

In the analysis plotted in **Extended Data Fig. 6i**, we altered the connections between mSPN and action/inaction encoding neurons on choice.

A modeled approach/avoid experiment is conducted by offering an option with reward = 1 arbitrary unit, cost = 1 arbitrary unit, and varying (10 values incremented from [0, 2] arbitrary units) physically proximity to another reward. An additional column is added to β_{direct} and $\beta_{indirect}$ in eq. (4) to reflect the fact that additional proximity to the other reward increases approach rate: $\beta_{direct} = \beta_{indirect} = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}$, where the upper row corresponds to approaching, the bottom row corresponds to not approaching, and the columns correspond to, from left to right, a reward-predominant decision-dimension, a cost-predominant decision-dimension, and a location-predominant decision-dimension. Reward is assigned a greater relative importance than cost or location (score of 3 arbitrary units versus 1 versus 1) in sSPNs, while cost is assigned a greater relative importance than reward or location (score of 3 arbitrary units versus 1 versus 1). daSNC activity is incremented by changing $z_{\mathrm{daSNC},i,P}$ in eq. (2) for all i and both pathways. Action and inaction values are calculated, and then choice is formed by averaging the results of 1000 Merton process simulations.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/dynamic_model_and_neural_net/direct_vs_indirect_pathway_proximity_theory.m.

Effects of sSPN, LHb, and daSNC activity on decision-space.

In **Figs. 5a,b**, b_{sSPN} (eq. (1)), z_{LHb} (eq. (5)), and $z_{\mathrm{daSNC},i,P}$ (for all i and a single pathway, eq. (2)) are incremented from 0 to 5 arbitrary units with 20 evenly spaced increments along each axis. Each of the 20x20x20 points are used to derive daSNC_i in eq. (2) and converted to decision-spaces via eq. (18). To limit the number of points in **Fig. 5a**, each point is shown with 0.5% probability.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/day_to_day_space_sampling/decision_space_by_circuit_activity.m.

Effect of decision-space on choice profiles.

In **Figs. 5c,f, Extended Data Figs. 7d,f**, we form decision-spaces using various decision-dimensions across incremented cortical reward and cost inputs, then classified the action values formed across those reward/cost inputs using a scoring system.

The scoring system, visualized in **Extended Data Fig. 7c**, is as follows:

Scores for “explore,” “riskiness,” “high action,” “exploit,” “safety,” “low action” are calculated by incrementing reward and cost on [-1 1] (arbitrary units) scales (9 increments are used for each of reward and cost in **Fig. 5c, Extended Data Fig. 7d,** 6 increments in **Figs. 5f, Extended Data**

Fig. 7f). The notation used here treats $v_{r,c,j}$ as the action value of the jth action at a certain reward and cost increment and $v_{r,c}$ as the set of those action values. In the plotted analysis, $k = 4$ actions are assigned action values.

- Explore. The tendency to pursue multiple actions simultaneously. Scored as the area of the region of reward and cost combinations with a Gini coefficient less than 0.25.

(26) $$\text{explore} = \sum_{\text{r}=-1}^{1} \sum_{\text{c}=-1}^{1} \left[\text{gini}\left(v_{r,c} \right) < 0.25 \right]$$

(27) $$\text{gini}\left(v_{r,c} \right) = \frac{\sum_{i=1}^{k} \sum_{j=1}^{k} |v_{r,c,i} - v_{r,c,j}|}{2k \sum_{j=1}^{k} v_{r,c,j}}$$

- Exploit. The tendency to pursue only one action. Scored at the area of the region of reward and cost combinations with a Gini coefficient greater than 0.5.

(28) $$\text{exploit} = \sum_{\text{r}=-1}^{1} \sum_{\text{c}=-1}^{1} \left[\text{gini}\left(v_{r,c} \right) > 0.5 \right]$$

- Riskiness. The combined value of actions when reward and cost are high. Scored by examining the combinations where both reward and cost are greater than 0.

(29) $$\text{riskiness} = \sum_{\text{r}=0}^{1} \sum_{\text{c}=0}^{1} \sum_{j=1}^{k} v_{r,c,j}$$

- Safety. The combined value of actions when reward and cost are low. Scored by examining the combinations where both reward and cost are less than 0.

(30) $$\text{safety} = \sum_{\text{r}=-1}^{0} \sum_{\text{c}=-1}^{0} \sum_{j=1}^{k} v_{r,c,j}$$

- High action. How often actions will have high action values. Scored as the area of the region of reward and cost combinations that have combined action value greater than 0.5.

(31) $$\text{high action} = \sum_{\text{r}=-1}^{1} \sum_{\text{c}=-1}^{1} \left[\sum_{j=1}^{k} v_{r,c,j} > 0.5 \right]$$

- Low action. How often actions will have low action values. Scored as the area of the region of reward and cost combinations that have combined action value less than 0.2.

(32) $$\text{high action} = \sum_{r=-1}^{1} \sum_{c=-1}^{1} \left[\sum_{j=1}^{k} v_{r,c,j} < 0.2 \right]$$

In the analyses plotted in **Fig. 5c** and **Extended Data Fig. 7d** and the examples in **Fig. 5d**, action value scores, as measured by the scoring definitions above, are compared when different decision-spaces are constructed but cortical input and system parameters are unchanged. The underlying action values across reward and cost levels resembles those from other analyses (see **Common parameters**) except for an addition of normal random noise (mean = 0, standard deviation = 1) to every element of β_{direct} (see eq. (4)).

The analysis in **Extended Data Fig. 7d** is similar, except for here, a weighted average is taken of action value scores, as measured by the scoring definitions above, between scenarios where different decision-spaces are constructed. A different $\mathbf{S}$ in eq. (3) is used depending on the specified dimensionality of direct pathway decision-space: $\mathbf{S} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$ for 1D, $\mathbf{S} = \begin{bmatrix} & & 0 \\ & \mathbf{I}_2 & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$ for 2D, $\mathbf{S} = \begin{bmatrix} & & 0 \\ & \mathbf{I}_3 & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$ for 3D, and $\mathbf{S} = \mathbf{I}_4$ for 4D. A weighted average of the five decision-spaces is calculated for three levels of sSPN activity (-1, 0, and 1).

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/day_to_day_space_sampling/subjective_value_scores_by_space.m.

In the analyses plotted in **Fig. 5f, Extended Data Fig. 7f**, for each of 1000 simulations, uncorrelated Gaussian white noise (mean = 0, standard deviation = 1) is added to every element of β_{direct} (see eq. (4)) and 6 by 6 grids of action values across reward and cost combinations are scored by the “explore,” “riskiness,” “high action,” “exploit,” “safety,” and “low action” metrics. Scores for each metric are compared across simulations and between decision-space groups. Observations that score in the top 10% by a metric are considered outliers. Outlier proportion is plotted in **Fig. 5f**. The means across the simulations of each score are plotted in **Extended Data Fig. 7f**.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/day_to_day_space_sampling/subjective_value_score_extremes.m.

Instance 2: sparse connectivity and feedforward.

In this section, we describe the instance of the model where cortex input to the system does not change over time and the activities of other circuit elements do not decay over time.

This instance of the model leads to a convenient formation of the model as a circuit with no time component. In this section, we frame this instance mathematically and then describe our related analysis examining the process of formation of the decision-space over time.

To focus on the portions of the circuit we analyze using this instance, we define here the subset of the circuit involving cortex, FSI, sSPN, daSNC, and mSPN.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/dynamic_model_and_neural_net/neural_network_model.m.

Cortical input.

A set of 4 cortical neurons, notated as C, is sampled at random from a population of 50 cortical neurons. Each neuron in C projects to one FSI and each of q SPNs, which each correspond to a decision-dimension. In our analysis, q is set to 4. This process is repeated 10,000 times per each pathway, forming 10,000 groups of 4 cortical neurons, 1 FSI, 4 dsSPNs (or isSPNs), and 4 dmSPNs (or imSPNs) for each pathway.

Defining FSI activity.

FSI activity is defined as a weighted sum of the activities of connected cortical neurons:

$$\mathrm{FSI}_C = \sum_{q \in C} w_{\mathrm{cortex} \to \mathrm{FSI}} \, \mathrm{cortex}_q + b_{\mathrm{FSI}} \tag{33}$$

where:

- C is a randomly sampled subset of cortical neurons.
- FSI_C is the activity of the FSI which receives projection from the cortical neurons in C (activity arb. u.)
- $w_{\mathrm{cortex} \to \mathrm{FSI}}$ is the connection weight between cortical neurons and FSIs (dimensionless)
- cortex_q is the activity of cortical neuron q (activity arb. u.)
- b_{FSI} affects the relative activity of all sSPN neurons (activity arb. u.)

In the current instance of the model, there are 10,000 FSIs that project to each of dSPNs and iSPNs. Each cortical neuron in C projects to FSI_C.

Defining sSPN activity.

sSPN activity is defined as a weighted sum of the activities of connected cortical neurons, divided by a weighted sum of the activities of connected FSIs, plus an additive shift b_{sSPN} applied to all sSPNs:

(10) $$\mathrm{sSPN}_{s,C} = \frac{1}{|C|} \sum_{q \in C} \frac{w_{q \to s} \, \mathrm{cortex}_q}{\mathrm{FSI}_C} + b_{\mathrm{sSPN}}$$

(copied from **Extended Data Fig. 4** for convenience)

where:

- C is a randomly sampled subset of cortical neurons.
- $\mathrm{sSPN}_{s,C}$ is the activity of an sSPN s that receives projection from cortical neurons in C. (activity arb. u.)
- $w_{q \to s}$ is the connection weight between cortical neuron q and sSPN s. The weight is equivalent to one of the first four principal components of the cortical activity of the four connected cortical neurons. sSPNs are separated into equal populations that correspond to the first, second, third, or fourth principal component. (dimensionless)
- FSI_C is the activity of the FSI which receives projection from the cortical neurons in C (activity arb. u.)
- cortex_q is the activity of cortical neuron q (activity arb. u.)
- b_{sSPN} represents the relative activity of all sSPN neurons (activity arb. u.)

In the current instance of the model, there are 40,000 neurons for each of dsSPNs and isSPNs. Activities are defined based on a feedforward network, so the simplification is made that sSPN activities are not affected by daSNC activities. All cortical neurons in C project to $\mathrm{sSPN}_{s,C}$ for all s, and similarly, FSI_C projects to $\mathrm{sSPN}_{s,C}$ for all s.

Defining daSNC activity.

The activity of the daSNC element corresponding to decision-dimension i and pathway P is defined as the weight from sSPNs corresponding to decision-dimension i and pathway P:

(11) $$\mathrm{daSNC}_{i,P} = \frac{1}{1+\exp\left(\frac{1}{n_{\mathrm{sSPN}}}\sum_{s\in i,P} w_{s\to \mathrm{daSNC},i,P}\cdot \mathrm{sSPN}_s + \mathrm{RMTg} - z_{\mathrm{daSNC},i,P}\right)}$$

(copied from **Extended Data Fig. 4** for convenience)

where:

- $\mathrm{daSNC}_{i,p}$ is the activity the daSNC neuron corresponding to decision-dimension i and pathway P. (activity arb. u.)
- n_{sSPN} is the count of sSPNs in each of the direct/indirect pathways
- $w_{s\to \mathrm{daSNC},i,P}$ is the connection weight from sSPN s to the daSNC neuron corresponding to decision-dimension i and pathway P. The weight is fixed in this instance of the model. (dimensionless)
- sSPN_s is the activity of SPN s (activity arb. u.)
- RMTg is RMTg activity. (activity arb. u.)
- $z_{\mathrm{daSNC},i,P}$ is the bias in the activity of a daSNC neuron corresponding to decision-dimension i and pathway P. (activity arb. u.)

In the current instance of the model, there are q daSNC neurons that receive projection from the 10,000 dsSPNs corresponding to each decision-dimension, and likewise q daSNC neurons that receive projection from the 10,000 isSPNs corresponding to each decision-dimension. RMTg also projects to all daSNC neurons.

In our analysis using this instance of the model, we arbitrarily set $w_{\mathrm{sSPN}\to \mathrm{daSNC},i} = 1$ arbitrary unit for both pathways and $z_{\mathrm{daSNC},i} = -5$ arbitrary units for all i for the direct pathway, and $z_{\mathrm{daSNC},i} = 5$ arbitrary units for all i for the indirect pathway. For simplicity, RMTg is set to 0.

Defining mSPN activity and decision-space.

mSPN activity is defined as a weighted sum of the activities of connected cortical neurons, divided by a weighted sum of the activities of connected FSIs. Here, unlike in the definition of sSPN activity in eq. (10), a $d_{i,P}$ term is multiplied to incorporate the weighting of mSPNs by dopamine:

(12) $$\mathrm{mSPN}_{m,C} = \frac{d_{i,P}}{|C|}\sum_{q\in C}\frac{w_{q\to m}\,\mathrm{cortex}_q}{\mathrm{FSI}_C}$$

(copied from **Extended Data Fig. 4** for convenience)

where:

- C is a randomly sampled subset of cortical neurons.
- $\mathrm{mSPN}_{m,C}$ is the activity of mSPN m that receives projection from cortical neurons in C. (activity arb. u.)
- $d_{i,P}$, which takes the value 0 or 1, is dopamine signaling to mSPNs corresponding to decision-dimension i and pathway P. $d_{i,P}$ is the realization of probabilistic weighting of decision-dimensions based on daSNC activity (see ***Conceptual Model***). (dimensionless)
- $w_{q \to m}$ is the connection weight between cortical neuron q and mSPN m. The weight is equivalent to one of the first four principal components of the cortical activity of the four connected cortical neurons. As described in ***Conceptual Model,*** sSPNs are separated into equal populations that correspond to the first, second, third, or fourth principal component. (dimensionless)
- FSI_C is the activity of the FSI which receives projection from the cortical neurons in C (activity arb. u.)
- cortex_q is the activity of cortical neuron q (activity arb. u.)

In the current instance of the model, there are 40,000 neurons for each of dmSPNs and imSPNs All cortical neurons in C project to $\mathrm{mSPN}_{m,C}$ for all m, and similarly, FSI_C projects to $\mathrm{mSPN}_{m,C}$ for all m.

Modeling SPN encoding of data, using Instance 2.

To explore the ability of SPNs to successfully encode data along decision-dimensions, even when cortex and SPNs are sparsely connected, we constructed networks with different degrees of dimensionality reduction (**Extended Data Fig. 4g**). A single pathway is considered. One type of network had 2 times dimensionality reduction (20 cortical neurons, 10 SPN), another had 10 times dimensionality reduction (100 cortical neurons, 10 SPN), and another had 100 times dimensionality reduction (1000 cortical neurons, 10 SPN), similar to what is found in the human brain[61,108].

For each type, we constructed modeled networks with cortex→SPN connections equal to principal components of cortical activity by simulating, for each analyzed pathway, a random symmetric positive definite matrix that is used as the cortical covariance matrix Σ_P (see **Defining sSPN activity**) via MATLAB's sprandsym() with density=1. Eigenvalues are arbitrarily specified as $\lambda_1 = 2$, $\lambda_2 = 2$, $\lambda_3 = 0.5$, $\lambda_4 = 0.2$, and $\lambda_5 = \lambda_6 = \ldots = \lambda_p = 0$, i.e. the first and second principal components are very important, the third somewhat important, the fourth slightly important, and the others unimportant. The weights $\mathbf{W}_P$ from cortical neurons to the q SPN circuit elements per pathway are then derived as the first q eigenvectors of Σ_P.

We incremented the number of cortical neurons that connected to each SPN from 2 to 10. The network was connected sparsely based on the specified number of connections from randomly

selected cortical neurons to each SPN. For simplicity in this analysis, we created a cortical signal that resembled the first principal component of cortical activity as a whole (regardless of connectivity to SPNs) and let the first cortical principal component to have large eigenvalue compared to the others (i.e. $\lambda_1 = 1$, $\lambda_2 = 0.1$, $\lambda_3 = \lambda_4 = \ldots = \lambda_p = 0$).

For each of the modeled networks (3 network types by 9 increments from 2 to 10), we simulated the process 1000 times. During each simulation, we calculated the ability of the network to discriminate between the large signal along the first cortical principal component and the absence of signal along the second cortical principal component, given its access to only a subset (2 to 10 cortical neurons) of the complete signal.

Then the SPNs encoding data along the first versus second decision-dimension were assessed in their ability to distinguish between signals along the first cortical principal component versus the second. This was quantified using the Bhattacharyya distance of the activity among SPNs encoding data along the first decision-dimension versus the second, assuming the subpopulations have mean activities μ_1 and μ_2 and standard deviations σ_1 and σ_2, respectively.

(34) $$D_B = \frac{1}{4}\frac{\left(\mu_1 - \mu_2\right)^2}{{\sigma_1}^2 + {\sigma_2}^2} + \frac{1}{2}\ln\left(\frac{{\sigma_1}^2 + {\sigma_2}^2}{2\sigma_1\sigma_2}\right)$$

where:

- μ_1 is the mean activity of activities in the first subpopulation (activity arb. u.)
- μ_2 is the mean activity of activities in the second subpopulation (activity arb. u.)
- σ_1 is the standard deviation of activities in the first subpopulation (activity arb. u.)
- σ_2 is the standard deviation of activities in the second subpopulation (activity arb. u.)

For the current analysis, a_{FSI} is arbitrarily set to 1 arbitrary unit and b_{FSI} is arbitrarily set to 0.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/dynamic_model_and_neural_net/dimension_discrimination_vs_sparsity.m.

Instance 3: full connectivity and dynamics.

In this section, we describe the instance of the model where each cortical neuron projects to each FSI, each FSI projects to each SPN (for dsSPN, isSPN, dmSPN, imSPN), and each cortical neuron projects to each SPN.

This instance of the model leads to a convenient formation of the model as a circuit of fewer elements (one FSI, one dsSPN, isSPN, dmSPN, and imSPN per the four decision-dimensions),

with a time component. In this section, we frame this instance mathematically and then describe our related analysis.

To focus on the portions of the circuit we analyze using this instance, we define here the subset of the circuit involving cortex, sSPN, daSNC, and mSPN.

Defining SPN and daSNC activity.

Here, the activities of SPN and daSNC elements and the weights from sSPN to daSNC are represented as a system of differential equations. Because FSI activity is not measured in the related analyses, cortical activity to pathway P after FSI normalization $x_{i,P}(t)$ is used as input to the system in the equations below.

For a model diagram, see **Fig. 4a**. Note that daSNC activity of 0 (i.e. average activity) leads to 0 change in SPN activity (due to the $1/2$ terms in eqs. (13) and (14).

(13) $$\tau \cdot \frac{\mathrm{d}s_{\mathrm{sSPN},i,P}(t)}{\mathrm{d}t} = -s_{\mathrm{sSPN},i,P}(t) + x_{i,P}(t) - w_{\mathrm{daSNC}\to\mathrm{sSPN},i,P} \cdot \left(y_{\mathrm{sSPN},i,P}(t) - \frac{1}{2}\right)$$

(copied from **Extended Data Fig. 6** for convenience)

(14) $$\tau \cdot \frac{\mathrm{d}s_{\mathrm{mSPN},i,P}(t)}{\mathrm{d}t} = -s_{\mathrm{mSPN},i,P}(t) + x_{i,P}(t) + w_{\mathrm{daSNC}\to\mathrm{mSPN},i,P} \cdot \left(y_{\mathrm{sSPN},i,P}(t) - \frac{1}{2}\right)$$

(copied from **Extended Data Fig. 6** for convenience)

(15) $$\frac{\mathrm{d}}{\mathrm{d}t} w_{\mathrm{sSPN}\to\mathrm{daSNC},i,P}(t) = \kappa \cdot s_{\mathrm{sSPN},i,P}(t)$$

(copied from **Extended Data Fig. 6** for convenience)

where:

(16) $$y_{\mathrm{sSPN},i,P}(t) = \frac{1}{1+\exp\left(w_{\mathrm{sSPN}\to\mathrm{daSNC},i,P}(t) \cdot s_{\mathrm{sSPN},i,P}(t) + \mathrm{RMTg} - z_{\mathrm{daSNC},i,P}\right)}$$

(copied from **Extended Data Fig. 6** for convenience)

- τ is the time constant related to the decay rate of activity (dimensionless)
- $s_{i,P}(t)$ is the SPN activity (either dsSPN, isSPN, dmSPN, or imSPN) corresponding to decision-dimension i and pathway P, as a function of time. (activity arb. u.)
- t is time (seconds)
- $x_{i,P}(t)$ is the cortical activity input, after FSI normalization, to an SPN corresponding to decision-dimension i and pathway P (activity arb. u.)
- $w_{\text{daSNC}\rightarrow\text{sSPN},i,P}$ is the connection weight from a daSNC neuron corresponding to decision-dimension i and pathway P to an sSPN (either dsSPN or isSPN) corresponding to decision-dimension i and pathway P. (dimensionless)
- $y_{i,P}(t)$ is the activity of the daSNC neuron corresponding to decision-dimension i and pathway P, as a function of time. (activity arb. u.)
- $w_{\text{sSPN}\rightarrow\text{daSNC},i,P}(t)$ is the connection weight from an sSPN (either dsSPN or isSPN) to a daSNC neuron corresponding to decision-dimension i and pathway P, as a function of time. (dimensionless)
- κ is a coefficient that determines the rate at which the connection from sSPNs to daSNC neurons change depending on sSPN (dsSPN or isSPN) activity. (dimensionless)
- $z_{\text{daSNC},i,P}$ is the bias in the activity of a daSNC neuron corresponding to decision-dimension i and pathway P. (activity arb. u.)
- RMTg is the output of RMTg, per eq. (5). (activity arb. u.)

In the current instance of the model, there are q dsSPNs, q isSPNs, q dmSPNs, q imSPNs, q daSNC neurons that each receive projection from a dsSPN, and q daSNC neurons that each receive projection from an isSPN. daSNC neurons corresponding to decision-dimension i and pathway P project to sSPNs and mSPNs of the same i and P.

Defining decision-space.

A decision-dimension is used to form decision-space at times when daSNC activity corresponding to the decision-dimension exceeds a threshold:

(17) $$\mathrm{S}_{i,P}(t)=\begin{cases}0 & y_{i,P}(t) < \text{threshold}\\ 1 & y_{i,P}(t) \geq \text{threshold}\end{cases}$$

(copied from **Extended Data Fig. 6** for convenience)

where:

- $y_{i,P}(t)$ is the activity of the daSNC neuron corresponding to decision-dimension i and pathway P, as a function of time. (activity arb. u.)
- $\mathrm{S}_{i,P}(t)$ is the application of dopamine release (via daSNC activity) to mSPN activity corresponding to decision-dimension i and pathway P. The notation here is used to match the notation in ***Instance 1***; $\mathrm{S}_{i,P}(t)$ is the (i,i) element of the diagonal matrix $\mathbf{S}_P(t)$ whose elements correspond to the weights assigned to the decision-dimensions, similar to in eq. (3). (dimensionless)

Defining action value.

Action value is defined here like in ***Instance 1****,* except mSPN activity (for dmSPN and imSPN) is defined as a function of time:

(35) $$v_{j,P} = \frac{1}{1+\exp\left(-\boldsymbol{\beta}_{j,P}\,\mathbf{s}_{\mathrm{mSPN},P}(t)-\alpha_{j,P}\right)}$$

where:

- $\boldsymbol{\beta}_{j,P}$ is a matrix of weights from dmSPNs to downstream action value encoding neurons for the direct pathway, or imSPNs to downstream inaction value encoding neurons for the indirect pathway. (dimensionless)
- $\mathbf{s}_{\mathrm{sSPN},P}$ is the activity of sSPNs corresponding to decision-dimension i and pathway P (activity arb. u.)
- t is time (seconds)
- $\alpha_{j,P}$ is an additive shift corresponding to the neuron encoding action j for the direct pathway or inaction j for the indirect pathway. (activity arb. u.)

As in the other instances of the model, there is one neuron encoding each $v_{j,P}$. So, there are k neurons encoding action values and k neurons encoding inaction values. Each of these neurons receives projection from all mSPNs of the corresponding pathway.

Modeling time-variant input, using Instance 3.

In **Figs. 4d-i,** simulated responses of dsSPN, isSPN, dmSPN, and imSPN are derived using the forward Euler method with step size 0.001s.

In **Figs. 4d,e,** the cortical input to sSPN elements corresponding to each of four example direct pathway decision-dimensions is represented by a vector of length 5001 (corresponding to 0s to

5s with increments 0.001s). Four arbitrary input vectors are used, one corresponding to each plotted decision-dimension:

(36) $2+\sin(t)\ ,\ 1+\cos(t)\ ,\ \sin(2t)\ ,\ \cos(2t)$

The first vector corresponds to a reward-predominant dimension (shown in green in **Fig. 4i**). A relatively large positive average value (2 arbitrary units) is assigned to it as an example of an important decision-dimension to a decision. A cost-predominant decision-dimension (second vector) is specified to be important, but less so, and novelty-predominant and location-predominant decision-dimensions (third and fourth vectors) are assigned to be relatively unimportant. The number of decision-dimensions used to form decision-space is averaged across time steps in the 5s simulation. In **d**, simulations are run across 100 evenly spaced increments of an addition to each vector at all timesteps. In **e,** simulations are run across 100 evenly spaced increments of $z_{\mathrm{daSNC},i}$ (for each pathway, depending on the simulation) in eq. (16) from -1 to 1 arbitrary units.

Figs. 4f,g show examples of the response of dsSPN, isSPN, dmSPN, and imSPN elements to different cortical inputs. In **Fig. 4f**, a cortical input of 10 arbitrary units for 2.5s is followed by an input of 20 arbitrary units for 2.5s. In **Fig. 4g**, a cortical input of 10 arbitrary units for 2.5s is followed by an input of 0 for 2.5s.

In the analysis shown in **Fig. 4h,** cortical inputs of 10 arbitrary units for 2.5s are followed by cortical inputs with prediction errors incremented by 0.1 from -1 to arbitrary units. These prediction errors are relative to the original cortical input of 10 arbitrary units. For example, for the prediction error of -1, there is a signal of 0 for 2.5s, and for the prediction error of 1, there is a signal of 20 for 2.5s. To find the change in the activities of circuit elements, their activities at 2.5s are subtracted from their activities at 5s.

Fig. 4i shows simulations for an example input with $\kappa=0$ in eq. (16) versus $\kappa=0.1$ arbitrary units. The cortical input is as follows: in decision-dimension 1 (e.g. reward-predominant), a cortical signal of 10 arbitrary units for the 0-1.23s timeframe and elsewhere a signal of 0; in decision-dimension 2 (e.g. cost-predominant), a cortical signal of 10 arbitrary units for the 1.25-2.43s timeframe and elsewhere a signal of 0; in decision-dimension 3, a cortical signal of 10 arbitrary units for the 2.5-3.73s timeframe and elsewhere a signal of 0; and in decision-dimension 4, a cortical signal of 10 arbitrary units for the 3.75-5s timeframe and elsewhere a signal of 0.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/dynamic_model_and_neural_net/sSPN_DA_mSPN_dynamic_interaction.m.

Parameters are set to common values, chosen arbitrarily, with physiologically accurate signs:

- in eqs. (13), (14): $\tau = 1$
- in eq. (13): $w_{\text{daSNC}\rightarrow\text{sSPN, direct}} = -1$
- in eq. (13): $w_{\text{daSNC}\rightarrow\text{sSPN, indirect}} = 1$
- in eq (14): $w_{\text{daSNC}\rightarrow\text{mSPN, direct}} = 1$
- in eq. (14): $w_{\text{daSNC}\rightarrow\text{mSPN, indirect}} = -1$
- in eq. (15): $\kappa = -0.01$
- in eq. (16): $\text{RMTg} = 0$
- in eq. (16): $z_{\text{daSNC}, i, P} = 0$ for all i and P
- in eq. (17): $\text{threshold} = 0.5$

Additionally, the following initial conditions, also chosen arbitrarily, are used across analyses:

- $s_{\text{sSPN}, i, \text{direct}}(0) = s_{\text{sSPN}, i, \text{indirect}}(0) = s_{\text{mSPN}, i, \text{direct}}(0) = s_{\text{mSPN}, i, \text{indirect}}(0) = 0$
- $w_{\text{sSPN}\rightarrow\text{SNc}, i, \text{direct}}(0) = w_{\text{sSPN}\rightarrow\text{SNc}, i, \text{indirect}}(0) = 1$

Movement of Circuit Activity Across Multiple Trials.

Here, we model changes in circuit activity between trials. We begin by forming advantage and cost functions that guide the realignment of the circuit. Using these, we explore how vulnerability versus resilience in disorder formation could be interpreted through the lens of the model.

Defining advantage and cost of circuit activity.

Advantage is defined here as the ability of the circuit to produce beneficial decision-spaces at a certain activity. The goal of the sSPN-GPi-LHb-RMTg-daSNC circuit in the model is to produce preferred decision-spaces for action valuation (**Figs. 6a-d**). For instance, in a laboratory environment when an animal routinely makes a choice to approach depending on reward level, a one-dimensional direct pathway decision-space with a reward dimension may be helpful. During a decision it may make sense for this animal to reach a circuit activity where forming a one-dimensional direct pathway decision-space is probable.

We represent this logic mathematically as a function of an n-element circuit $\left\{X_1, X_2, \ldots, X_n\right\}$. In our analysis, we focus on either FSI and sSPN, holding the rest of the circuit elements fixed at default values (see **Common parameters**); or sSPN, LHb, and daSNC, holding the rest of the circuit elements fixed at default values. The advantage of a certain circuit activity is defined as a weighted sum of probabilities the direct pathway decision-space occurs and the benefit of forming each decision-space:

(19) $$\text{advantage}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \sum_{l=1}^{2^q} \text{score}_l \cdot \text{P}(\text{space}_l | (X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n))$$

(copied from **Extended Data Fig. 8** for convenience)

where:

- $\{X_1, X_2, \ldots, X_n\}$ are elements of the circuit with activities $x_1, x_2, \ldots, x_n$. (activity arb. u.)
- q is the count of decision-dimensions
- score_l is a coefficient corresponding to the preference for a given direct pathway decision-space. (dimensionless)

The probability each decision-space forms is derived from probability its decision-dimensions individually are used during the decision (daSNC_i in eq. (2), here notated as d_i):

(37) $$\text{probability of formation of a decision-space } l = d_1\left(1-d_1\right) d_2\left(1-d_2\right) \cdot \ldots \cdot d_m\left(1-d_m\right)$$

The rationale for this formation of advantage is theoretical. Elsewhere, we show that direct pathway decision-spaces of different dimensionality are beneficial (and may be used by rodents) for tasks of different difficulties (**Fig. 2**). We also show that certain decision-spaces are beneficial with certain levels of cortical noise (**Extended Data Fig. 5**) and for obtaining different types of action values (**Figs. 5c-f**). We represent this as an assignment of greater value to certain decision-spaces, given external and internal contexts and the task at hand.

Cost, here, is defined as the difference between the circuit activity and a baseline activity. For most scenarios, the circuit might be best served searching for the circuit activity with the highest advantage. However, there is an obvious counterexample: it could be that it is easiest to form preferred decision-spaces at extremely unusual circuit activity (e.g. very high sSPN, very high LHb, very high daSNC), and only slightly more difficult to form that decision-space at closer to average circuit activity (average sSPN, low LHb, average daSNC). It may be more advantageous for the circuit to shift to the latter activity.

Therefore, we introduce a cost function to form *net advantage*. We then use net advantage to define the circuit activities that are the most beneficial. The concept of baseline circuit activity is introduced here in order to define cost. This can be interpreted as the circuit activity outside of decision-making.

(20) $$\text{cost}(X_1 = x_1,\, X_2 = x_2, \ldots, X_n = x_n) = \left\| \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^{\mathsf{T}} - \begin{bmatrix} x_{1,\text{baseline}} & x_{2,\text{baseline}} & \cdots & x_{n,\text{baseline}} \end{bmatrix}^{\mathsf{T}} \right\|_2$$

(copied from **Extended Data Fig. 8** for convenience)

where:

- $\{X_1, X_2, \ldots, X_n\}$ are elements of the circuit with activities $x_1, x_2, \ldots, x_n$. (activity arb. u.)
- Outside of decision-making, $\{X_1, X_2, \ldots, X_n\}$ have baseline activities $x_{1,\text{baseline}}, x_{2,\text{baseline}}, \ldots, x_{n,\text{baseline}}$. (activity arb. u.)

Net advantage is defined as advantage minus cost multiplied by a constant:

(21) $$\text{net advantage}(X_1 = x_1, \ldots, X_n = x_n) = \text{advantage}(X_1 = x_1, \ldots, X_n = x_n) - \text{constant} \cdot \text{cost}(X_1 = x_1, \ldots, X_n = x_n)$$

(copied from **Extended Data Fig. 8** for convenience)

where:

- Functions for reward and cost are taken from eqs. (19) and (20), respectively.
- constant alters the weight given to cost compared to reward. (dimensionless)

Visualizing advantage, cost, and net advantage.

Fig. 6b shows an example of how a circuit forms advantage per eq. (19). The advantage scores score_i are randomly generated (normal distribution, mean 0, standard deviation 1) for each of the 16 possible decision-spaces formed from four decision-dimensions. sSPN and daSNC activity are incremented across a 10x10 grid of sSPN and daSNC activities which each range from 0 to 10 arbitrary units. The activities of other circuit elements are set to default values in ***Instance 1*** of the model (see **Common parameters**).

Cost in **Fig. 6c** is formed using eq. (20). Distance from the circuit baseline point is measured as Euclidean distance in the two plotted decision-dimensions.

Net advantage in **Fig. 6d** is calculated per eq. (21). For this example, the cost coefficient $\text{constant} = 1$ (dimensionless coefficient).

Fig. 6f similarly shows net advantage but after incrementing sSPN, LHb, and daSNC activities. Similar to in the example in **Fig. 6d**, a set of scores are independently randomly generated and constant in eq. (21) is set arbitrarily to 1 (dimensionless coefficient).

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/circuit_trajectories/advantage_cost_net_advantage_example.m.

Defining direction of circuit movement.

The direction of movement is defined as a search for the most optimal circuit activity to produce advantageous direct pathway decision-spaces. To understand how a circuit governed by eq. (21) might adjust over the course of multiple trials, we relate the adjustment of circuit activity to net advantage. We specify that the circuit adjusts so that it can more easily reach advantageous decision-spaces. This constitutes a movement in the baseline activity of eq. (20).

Thus, the circuit adapts in the direction of the gradient of net advantage.

(22) $$\frac{\Delta \left[x_{1,\text{baseline}} \;\; x_{2,\text{baseline}} \;\; \cdots \;\; x_{n,\text{baseline}} \right]^{\intercal}}{\text{trial}} = \text{rate} \cdot \nabla \text{ net advantage}(X_1 = x_1, \ldots, X_n = x_n)$$

(copied from **Extended Data Fig. S8** for convenience)

where:

- $\left\{ X_1, X_2, \ldots, X_n \right\}$ are elements of the circuit with activities $x_1, x_2, \ldots, x_n$. (activity arb. u.)
- Outside of decision-making, $\left\{ X_1, X_2, \ldots, X_n \right\}$ have baseline activities $x_{1,\text{baseline}}, x_{2,\text{baseline}}, \ldots, x_{n,\text{baseline}}$. (activity arb. u.)
- rate is a coefficient that affects the speed of movement of the circuit per trial. (trial^{-1})

Effect of initial circuit activity on future trials.

In **Figs. 6e,g, Extended Data Fig. 8a**, several examples illustrate trajectories of circuit movement, as defined in eq. (21).

In these analyses, we assume that decision-dimensions are assigned equal importance by sSPN, i.e. (1) in eq. $b_{\text{sSPN}} = 0$. The probability that an individual decision-dimension is used to form decision-space is calculated using eq. (2). Advantage scores score_i are randomly generated (normal distribution, mean = 0, standard deviation = 1 arbitrary unit) for each of the 16 possible decision-spaces created from four dimensions. The value of constant in eq. (21) is set to 1 arbitrary unit. For each increment of a 15x15 grid of sSPN and FSI (**Fig. 6e**) or 15x15x15 grid of sSPN, LHb, and daSNC (**Fig. 6g, Extended Data Fig. 8a**), net advantage is calculated for the scenario where that activity is the circuit baseline, i.e. $X_1 = x_{1,\text{baseline}}, X_2 = x_{2,\text{baseline}}, \ldots, X_n = x_{n,\text{baseline}}$.

The gradient of this grid is approximated via MATLAB's gradient() routine. Trajectories are integrated from the gradient via MATLAB's streamline(), which uses the forward Euler method.

In the upper panel of **Fig. 6e**, circuit baseline points are shown for each increment of Euler's method. The bottom panel shows the trajectories of 20 randomly selected points on the [0, 5] arbitrary units range for sSPN and FSI.

In **Fig. 8g**, starting points are selected to form a 5 by 5 by 5 grid incremented along sSPN, LHb, and daSNC axes. At each point in a trajectory line, the probability a dimension is used to form decision-space is calculated via eq. (2). Then, during plotting, these calculated probabilities are interpolated using MATLAB's patch() routine.

In **Extended Data Fig. 8a**, the gradient is plotted using MATLAB's streamslice(). Then the trajectories of five randomly selected are plotted.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/circuit_trajectories/sSPN_DA_LH_trajectories.m and https://github.com/dirkbeck/DM_space_model/blob/main/circuit_trajectories/sSNC_DA_LH_trajectories_examples2.m

Effect of altered advantage score on future trials.

In **Extended Data Fig. 8b,c**, we examine the effect of a change to advantage score_l in eq. (19) as a modeled task is being performed.

The circuit adjusts over 100 trials. At the beginning, advantage score is set to 1 (dimensionless coefficient) for the non-D decision-space and 0 for every other decision-space. Over the first 20 and the last 10 time-steps (between the dashed lines on the plots), advantage score is incremented for the 4D decision-space but not others. Over time steps 21-90, advantage score is incremented for the non-D decision-space.

In the plot, the value of each decision-space is divided by the total value of all decision-spaces to form a ratio.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/circuit_trajectories/changing_space_value_between_trials.m.

Rationale for the computational framework.

During the formation of our computational model, we considered several alternative modeling techniques, including neural networks, Hidden Markov Models, Bayesian methods, State Space Models, biophysics models, drift diffusion models, and others.

Here, we outline justifications for the modeling methods we chose and suggest other modeling techniques that may achieve similar results.

Modeling choices: dimensionality reduction.

We model the decision-making axes used to make choices (decision-dimensions) as the principal components of cortical data to link the processes of the circuit to operations commonly performed in data science. The decision-space model could easily be modified in a way that would preserve its concept, for example by replacing the principal components with the independent components[114] of cortical data. However, modifications may come at the cost of computational convenience or interpretability.

Modeling choices: choice from physiology.

We sought to derive choice and deliberation time of an arbitrary number of potential actions. Drift diffusion models have been successful in other work involving the Basal Ganglia[115], but classically only model two possible outcomes, although extensions have been made for the multi-outcome case. We designed our model to have a similar framework to drift diffusion models but be suited to an unlimited number of possible actions. Our model successfully reproduces choice in the T-maze task (**Extended Data Figs. 2b,c**). Our goal could also be achieved through models of interacting populations[116], though in our objective, we are more considered with choice outcomes (choice, deliberation time, deliberation time distribution) than neural recruitment to produce it. Alternatively, we might find success by treating choice as a Bayesian process arising from circuit activity[117] or through an intermediate layer such as in a Hidden Markov Model[118], though deliberation time is not commonly derived using these classes of models.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/Cross%20Correlation%20Pattern%20Counts/Neural%20Network/createHistogramsByTaskTypeOrConcentration.m (**Extended Data Fig. 2b**) and https://github.com/dirkbeck/DM_space_model/blob/main/Cross%20Correlation%20Pattern%20Counts/Neural%20Network/newDirkCode.m **Extended Data Fig. 2c**)

Modeling choices: physiological connections.

We modeled the overall activities of circuit elements using relative firing rates because it achieved our goal of linking the overall activities of circuit brain regions. While models that model the timing of neuron spikes have been used to precisely describe connections in portions of the circuit[119,120], the decision-space conceptual model does not require their added granularity.

Reasoning behind the FSI model.

In eq. (1), scaling of cortical data serves the important functional role of normalizing $\mathbf{x}_P$ so that the circuit can make sense of data coded across the wide range of firing rates that are observed in the cortex[121] (**Extended Data Fig. 1a**). Previous anatomical work and physiological analysis has shown that many cortical neurons synapse to one FSI[121,122]. The cortical to FSI connection is excitatory[123]. This would suggest that FSI surveys the cumulative activity of cortex. Strengthening this hypothesis, physiological analysis of connected FSI and cortical neurons in the current work shows that FSI activity scales linearly with cortical activity[10,124] (see **Extended Data Fig. 3k,l**).

In the model, the parameter a_{FSI} is related to the strength of connection between cortex and FSI. The parameter b_{FSI} is related to firing rate of FSI when $\mathbf{x}_P = \mathbf{0}$.

For the FSI and SPN relationship, anatomical and physiological work has shown an inhibitory effect of FSI on SPN[125]. The algebraic operator that best represents this inhibition is less clear: it could be thought of as a subtraction or as a division depending on the experimental evidence used to model[126,127].

A subtraction possibility:

(38) $$\mathbf{s}_{\mathrm{sSPN}}\left(a_{\mathrm{FSI}}\right) = \mathbf{x}_P - a_{\mathrm{FSI}}$$

A division possibility:

(39) $$\mathbf{s}_{\mathrm{sSPN}}\left(a_{\mathrm{FSI}}\right) = \frac{\mathbf{x}_P}{a_{\mathrm{FSI}}}$$

where:

- a_{FSI} is the weight of cortex→FSI connection (dimensionless)
- $\mathbf{x}_P$ is the activities of the cortical neurons in a given pathway P (activity arb. u.)

The analytic relationship used here leads to a convenient interpretation from the perspective of data processing. It can also be thought of as approximating subtraction or division depending on the scale of a_{FSI}. Below, we show a series approximations of eq. (1) as a function of connection strength from cortex, with input from cortex and other parameters fixed.

Substitution in terms of a_{FSI}:

(40) $$\mathrm{s}_{\mathrm{sSPN}}\left(a_{\mathrm{FSI}}\right)=\frac{1}{a_{\mathrm{FSI}}\cdot\left\|\mathbf{x}_{P}\right\|_{2}+b_{\mathrm{FSI}}}\mathbf{x}_{P}\mathbf{W}+b_{\mathrm{sSPN}}$$

Truncated Taylor series at $a_{\mathrm{FSI}}=0$:

(41) $$\mathrm{s}_{\mathrm{sSPN}}\left(a_{\mathrm{FSI}}\right)=\mathrm{b}_{\mathrm{sSPN}}+\frac{\mathbf{x}_{P}\mathbf{W}}{b_{\mathrm{FSI}}}-\frac{\left\|\mathbf{x}_{P}\right\|_{2}\mathbf{x}_{P}\mathbf{W}}{{b_{\mathrm{FSI}}}^{2}}\cdot a_{\mathrm{FSI}}+\mathcal{O}\left({a_{\mathrm{FSI}}}^{2}\right)$$

Truncated Laurent series at $a_{\mathrm{FSI}}=\infty$:

(42) $$\mathrm{s}_{\mathrm{sSPN}}\left(a_{\mathrm{FSI}}\right)=b_{\mathrm{sSPN}}+\frac{\mathbf{x}_{P}\mathbf{W}}{\left\|\mathbf{x}_{P}\right\|_{2}}\cdot\frac{1}{a_{\mathrm{FSI}}}-\mathcal{O}\left(\frac{1}{{a_{\mathrm{FSI}}}^{2}}\right)$$

Thus, when FSI is weakly connected to cortex, the formula somewhat resembles a subtractive operation, and when FSI is strongly connected to cortex, the formula somewhat resembles a division operation. The ratio between the a_{FSI} and b_{FSI} parameters is important here (the first series, before truncation, converges when $\frac{\left|b_{\mathrm{FSI}}\right|}{a_{\mathrm{FSI}}}>\left\|\mathbf{x}_{P}\right\|_{2}$ and the second, before truncation, when $\frac{\left|b_{\mathrm{FSI}}\right|}{a_{\mathrm{FSI}}}<\left\|\mathbf{x}_{P}\right\|_{2}$).

Inferring decision-space from SPN activity and choice.

The decision-space can be inferred from environmental or experimental inputs in conjunction with decision-making data (**Extended Data Figs. 1q,r**). The method here requires that many decision-making experiments have been run with a similar apparatus but different parameters (for instance, light level), and that average sSPN activity has been measured during the decisions. The parameter data is stored in a matrix $\mathbf{X}\in\mathbb{R}^{n\times p}$, where n here is a separate trial with separate inputs (rather than different inputs across time steps, as in the description of ***Instance 1***) and p

is the number of features of the experiment that may be encoded by cortex, for instance temperature, music volume, or light level (similar to the description in ***Instance 1***).

The process involves three steps:

Labeling sessions.

Recorded decisions are labeled using attributes of the process by which the decision was made. This might be achieved in a rodent task that measures response to music volume, for instance, by clustering sessions based on heart rate and distance traveled.

Constraints based on SPN activity.

The sessions are split by label. $\mathbf{X}$ is standardized such that each column has a mean of 0 and standard deviation 1. Then a linear regression is run on each labeled subset $\mathbf{X}_l$ to find coefficients $\mathbf{b}_l$ that map those observations to predicted SPN activities (averaged across SPN decision-dimensions) $\hat{\mathbf{y}}_l$ for each session in the subset (here, the subscript l is used to indicate all elements of the subset):

(43) $$\hat{\mathbf{y}}_l = \mathbf{X}_l \mathbf{b}_l + b_0$$

where:

- $\hat{\mathbf{y}}_l$ is predicted SPN activities
- $\mathbf{X}_l$ is a subset of experimental parameter data (see *Labeled sessions)* corresponding to one label
- $\mathbf{b}_l$ and b_0 are the linear regression coefficients that map the experimental parameters (see **Labeled sessions***)* to sSPN activity

From the calculated $\mathbf{b}_l$, we can guess the principal axis dimensions that make up decision-space $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_q$ and the presence of each decision-dimension $e_{li} \in \{0, 1\}$ in each labeled subset. During this process, we consider $\mathbf{b}_l$ as the sum of the dimensions in decision-space in the subset:

(44) $$\mathbf{b}_l = \sum_{i=1}^{q} e_{li} \mathbf{w}_i$$

Using this framework, we can form a rule which we can use to assign labels to decision-dimensions and hypothesize $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_q$:

> If $\mathbf{b}_A$ has high correlation with $\mathbf{b}_B$, then one of $\mathbf{b}_A$ or $\mathbf{b}_B$ must correspond to a higher-dimensional decision-space that also includes the dimensions of the other.

For example, we might have data with labels A, B, C, and D and corresponding $\mathbf{b}_A, \mathbf{b}_B, \mathbf{b}_C, \mathbf{b}_D$, where $\mathbf{b}_A$ is correlated with $\mathbf{b}_D$, $\mathbf{b}_B$ is correlated with $\mathbf{b}_D$, and the other possible pairs uncorrelated. It follows from the rule above that $\mathbf{b}_D$ is at least a two-dimensional decision-space including dimensions from $\mathbf{b}_A$ and $\mathbf{b}_B$, and that $\mathbf{b}_A$ and $\mathbf{b}_B$ are at least one-dimensional decision-spaces. So, we would hypothesize that there are two SPN-encoded decision-dimensions used during the decision, $\mathbf{w}_1$ and $\mathbf{w}_2$, and that subsets A and D use $\mathbf{w}_1$, subsets B and D use $\mathbf{w}_2$, and subset C does not use either.

The logic to construct constraints I and II are as follows. When $e_{l1} = e_{l2} = \ldots = e_{lq} = 0$ (i.e. a non-D decision-space), $\hat{\mathbf{y}}_l = b_0$. The coefficients assigned to the non-D decision-space $\mathbf{b}_{l:\,\text{0D space}}$ should have low correlation with any of the axes $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_q$. Therefore, it is expected that we find one $\mathbf{b}_l$ that has low correlation with the other $\mathbf{b}_l$, and we can label this $\mathbf{b}_{l:\,\text{0D space}}$. Further, when e_{li} for one i is equal to 1 and for all other i is equal to 0 (i.e. a 1D decision-space), $\hat{\mathbf{y}}_l = \mathbf{w}_{i\text{ in space}} + b_0$. This decision-space will have high correlation with $\mathbf{w}_{i\text{ in space}}$ but low correlation with other $\mathbf{w}_i$. Therefore, it is expected that we find several $\mathbf{b}_{l:\,\dim i\text{ space}}$ that are not correlated with one another, or $\mathbf{b}_{l:\,\text{0D space}}$, but may be related through the multi-dimensional decision-spaces in C. Finally, when e_{li} is equal to 1 for multiple i, $\mathbf{b}_l$ is calculated as a sum per eq. (44). Therefore, it is expected that some $\mathbf{b}_{l:\,\dim i \,\cap\, \dim j\text{ space}}$ are linear combinations of $\mathbf{b}_{l:\,\dim i\text{ space}}$ and $\mathbf{b}_{l:\,\dim j\text{ space}}$.

The example shown in **Extended Data Fig. 1r** uses simulated data with 2 experimentally observed features (for instance, temperature and music volume) and 100 observations. $\mathbf{X}$ was constructed as a matrix of i.i.d. Gaussian variables with mean 0 and standard deviation 1. The ground-truth principal component matrix $\mathbf{W}$ was constructed as a 2x2 matrix of i.i.d. Gaussian variables with mean 0 and standard deviation 1. Then, for each of the 100 observations, a decision-space was randomly assigned in a reference dataset. Each decision-dimension for each observation was treated as an i.i.d. uniform variable, and if the variable corresponding to the observation and the decision-dimension exceeded a value (here, 0.5), then the dimension was considered as incorporated in decision-space. The randomly generated decision-spaces were each assigned their own label. Then a simulated SPN activity was created for each of the 100 observations was created by multiplying the ground-truth $\mathbf{W}$ by the decision-space used in each observation, similar to in eq. (1). i.i.d. Gaussian noise (mean 0, standard deviation 0.5) was then added to create simulated SPN activity observations. Using this simulated data, we ran a linear regression for each labeled subset via MATLAB's fitlm() routine. The linear regression fits are plotted as surfaces in **Extended Data Fig. 1r**. The slopes, with respect to temperature and music volume, were, for label A, -0.15 and 0.48, respectively; for label B, -437.90 and -9.72, respectively; for label C, -9.01 and -274.51 respectively; and for label D, -155.45 and -199.83. Slopes for label are relatively small, so it is assigned to "decision-space not formed" (matching the reference

dataset). Of the remaining labels, label D is closer to an additive combination of labels B and C than other permutations, so label D is assigned the "2D decision-space" (matching the reference dataset), while labels B and C are each considered 1D decision-spaces (matching the reference dataset).

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/model_overview/dimensionality_from_SPN_activity.m

Tests using choice.

We can then combine the constraints in Step 2, developed using SPN activity, with an analysis of choice given the hypothesized SPN-encoded decision-dimensions. In Step 2, a set of $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_q$ are hypothesized. Here, choice at different levels of those $\mathbf{w}_i$ are compared across the labeled subsets. The hypothesis in Step 2 is supported if choice, for each labeled subset, is correlated with the decision-dimensions hypothesized to be used to form decision-space but not correlated with the decision-dimensions hypothesized not to be used to form decision-space.

An example is shown in **Extended Data Fig. 1r.** For the case where decision-space is not formed, choices do not correlate with any hypothesized SPN-encoded decision-dimension. For a 1D decision-space, choices correlate with one hypothesized SPN-encoded decision-dimension. For a 2D decision-space, choices correlate with two hypothesized SPN-encoded decision-dimensions.

In the plotted examples, we plot simulated choices in each of the four decision-spaces used in the analysis. Choices are simulated from reference dataset (i.e. absent noise added during simulation) average sSPN activities by session y by treating $Z = \frac{1}{1+\exp(-y+\text{i.i.d. Gaussian noise})}$ as a random variable, where $Z<0.5$ corresponds to a "turn left" action and $Z \geq 0.5$ to a "turn right" action. The threshold of 0.5 is chosen because it represents the expected value of Z at average sSPN activity ($y=0$). In the plots, we interpolate possible subject values from choice at different combinations of the two decision-dimensions $\mathbf{w}_1$ and $\mathbf{w}_2$, as derived in Step 2. For each action for each labeled subset, a logistic regression is used to convert actions to value of action across the grid. As expected, observations of label A, assigned to "decision-space not formed" have little correlation with either of the putative decision-dimensions derived in Step 2; observations of labels B and C, assigned to 1D decision-spaces, are correlated with one putative decision-dimension but not the other; and observations of label D, assigned to the "2D decision-space," are correlated with both.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/model_overview/dimensionality_from_SPN_activity.m

Testing the Model Through Analysis of Neural Data.

As a further test of the decision-space model, we examined the relationships between behavior and neural activity in tasks that required different reward versus cost dimensions. The analysis, new to the current work, was performed using the Corticostriosomal Circuit Stress Experiment database (published with Friedman et al., 2017). We found more functionally connected sSPN and mSPN in tasks that were difficult. We then analyzed cortex, FSI, sSPN, and mSPN during these tasks and found evidence of dimensionality reduction from cortical neurons to SPNs.

Defining decision difficulty by task.

We defined decision difficulty through deliberation time, calculated as the time between when the door opened during the T-maze task and when the animal made a movement between one end of the maze or the other. Deliberation time distributions were analyzed for each trial group (for control and stress: cost-benefit cost, benefit-benefit, cost-cost, non-conflict cost-benefit). Skewness was calculated using MATLAB's skewness() routine. Example distributions are shown in **Extended Data Figs. 2b, 3a,b** (6 animals, 35 sessions) and summaries across groups in **Extended Data Figs. 2d, 3c,d** (14 rats, 249 sessions).

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/Cross%20Correlation%20Pattern%20Counts/Neural%20Network/createHistogramsByTaskTypeOrConcentration.m (**Extended Data Fig. 2b**) and https://github.com/dirkbeck/DM_space_model/blob/main/Cross%20Correlation%20Pattern%20Counts/Pattern%20Analysis/createSkewnessBarChartOfTRAndCBAcrossControlAndStress2.m (**Extended Data Fig. 2d**)

Fitting the modeled deliberation times.

To replicate the deliberation time distributions with a computational model, we leveraged a diffusion model where two functions would both start at zero and would continue towards either a positive or negative pre-defined threshold that represented either performing an action (positive threshold) or not performing an action (negative threshold). The first threshold that was reached by either function was which decision we considered as selected. The x-axis was the progression of time as the functions ran. The deliberation time for this model was the amount of time (in seconds) it took for the first function to reach its threshold. To match the experimental distribution times analyzed in **Extended Data Fig. 2b**, we ran the drift diffusion process adjusting drift rate, the thresholds, and noise until the modeled deliberation time distribution had a similar median and skew to experimental data from **Extended Data Fig. 2c** (https://doi.org/10.7910/DVN/SMKW0I). An additional 1.5 seconds was uniformly added to modeled distribution times.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/Cross%20Correlation%20Pattern%20Counts/Neural%20Network/createHistogramsByTaskTypeOrConcentration.m.

Connected SPNs through cross-correlation.

We used the Corticostriosomal Circuit Stress Experiment database to identify sSPN and mSPN among recorded neurons in the striatum. For details, see the Supplemental Materials & Methods of Friedman et al. (2017). 14785 cells across 14 control animals were analyzed. The total number of identified striosomal and matrix neuron per task: NCB = non-conflict cost-benefit (sSPNs = 14, mSPNs = 260), CC = cost-cost (sSPNs =46 , mSPNs = 400), CBC = cost-benefit conflict (sSPNs = 84, mSPNs =717), BB = benefit-benefit easy (sSPNs = 50, mSPNs =515 , chocolate milk concentration <50), BB = benefit-benefit difficult (sSPNs =33, mSPNs = 731, chocolate milk concentration >=50).

We binned the firing rates of sSPN and mSPN during the -3s to 3s window across all tasks into 1-5 bins. We used these firing rates to determine cross-correlation (MATLAB's xcorr()) between sSPN and mSPN recorded in the same session. Correlated pairs were defined as paired sSPN and mSPN that had a linear regression fit with correlation squared (MATLAB's corrcoef()) > 0.5 and significance $p<0.04$. To obtain the percentage of significantly correlated neurons, we counted the number of pairs that met the threshold for correlation and divided by the total number of identified pairs. Examples are plotted in **Extended Data Figs. 2e,f** and counts in **Extended Data Fig. 2g.**

To determine significance threshold for the counts of correlated pairs, we shuffled the sSPN and mSPN pairs across the database. We then performed the process above on the shuffled data. Using these shuffled pairs, we formed a distribution of correlations that might happen by chance. The threshold of significance was set to the 3 standard deviation mark among this distribution of shuffled pairs.

To produce **Extended Data Figs. 2e,f**, first follow the steps in run_me.m to generate 'pairsTableControl.mat', 'pairsTableStress.mat', and 'pairsTableStress2.mat' are first generated.

Then run the following lines of code in MATLAB:

examplePairedNeuronPLot('twdb_control', 'PLSvsFSI', 54); % Figure S2E
examplePairedNeuronPLot('twdb_stress', 'PLSvsFSI', 32); % Figure S2F

For the code that calls the figure creation in **Extended Data Fig. 2g**, see https://github.com/dirkbeck/DM_space_model/blob/main/Cross%20Correlation%20Pattern%20Counts/Fig7A_triplet-example/updated_cross_correlation.m

The fitting occurs in: https://github.com/dirkbeck/DM_space_model/blob/main/Cross%20Correlation%20Pattern%20Counts/Fig7A_triplet-example/matrix_strio_plot_dynamics.m

Connected SPNs through Granger causality.

sSPN and mSPN pairs were identified similarly to the previous section. 14785 cells across 14 control animals were analyzed and 25758 cells across 9 stress animals. The total identified number of striosomal and matrix cells for each of the tasks are the following: Control CC = cost-cost (sSPNs =46 , mSPNs = 400), Control BB = benefit-benefit (sSPNs =83 , mSPNs = 1246), Control CBC = cost-benefit conflict (sSPNs = 84, mSPNs =717), Stress CBC = cost-benefit conflict (sSPNs =41 , mSPNs =898), Stress BB = benefit-benefit (sSPNs = 156, mSPNs = 2813).

Then firing rates were determined over the span of the session by binning each trial into bins of 100ms. Granger causality (MATLAB's gctest()) was run on these sSPN and mSPN firing rate pairs to determine whether each pair was functionally connected. In control animals, 92 connected pairs were identified for the cost-benefit conflict task, 77 connected pairs for the cost-cost task, 1 connected pair for the non-conflict cost-benefit task, and 116 pairs for the benefit-benefit task. In stress animals, 110 connected pairs were identified for the cost-benefit conflict task and 292 connected pairs for the benefit-benefit task.

Next, for each of the trials with a functionally connected sSPN and mSPN pair, the class of motif was determined as either sSPN excited / mSPN excited, sSPN excited / mSPN inhibited, sSPN inhibited / mSPN excited, or sSPN inhibited / mSPN inhibited. These classes were formed by identifying whether, separately, the neurons were excited, inhibited, or neither in each of 5 blocks per each trial ([-15s, -3s], [-3s 0s], [0s 2.5s], [2.5s, 4.5s], [4.5s 20s]). A motif was counted if in a certain block each neuron was either excited or inhibited.

Excited or inhibited classifications were determined through inter-spike interval analysis, plotted in **Extended Data Figs. 2h-j**. The median inter-spike interval for each neuron across the full trial was calculated. A neuron was considered inhibited during the periods of the trial when inter-spike exceeded median. Conversely, a neuron was considered inhibited during periods when inter-spike interval fell below median. Excited blocks were defined as blocks where excitation time exceeded inhibition time inhibited by 10%. Inhibited blocks were defined, oppositely, as blocks where inhibition time exceeded excitation time by 10%. Blocks that reached neither threshold were not classified.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/Cross%20Correlation%20Pattern%20Counts/Pattern%20Analysis/plotBins.m.

Patterns were counted across the 5 bins and added across trials within the same group (e.g. control cost-benefit conflict). The total count, as plotted in **Fig. 2e**, was divided by the number of trials to form an average number of functionally connected neurons across the 5 bins of the trial.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/Cross%20Correlation%20Pattern%20Counts/Pattern%20Analysis/allMaps/Maps%20By%20Task%20Type/createLinearPatternCountGraphsByTaskType.m.

Significance was determined similarly to the section above. Pairs of neurons, not necessarily functionally connected, were shuffled. From a shuffled distribution, the 3 standard deviation mark was identified and considered to be the threshold of significance.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/Cross%20Correlation%20Pattern%20Counts/Pattern%20Analysis/createRandomPatterns.m

Analyzing neural dimensionality reduction.

In the analysis plotted in **Fig. 2f,** using data from the Corticostriosomal Circuit Stress Experiment database, we analyzed firing rates of 1) PL neurons that project to striatum (i.e. sSPN-projecting cortex, 221 sessions, 2-47 neurons per session), 2) FSIs (96 sessions, 2-12 neurons per session), 3) sSPNs (27 sessions, 2-9 neurons per session), and 4) mSPN (13 sessions, 2-6 neurons per session). For methods of classification see Friedman et al. (2017). Spike data was converted to firing rates by separating the spikes into 5-10 bins.

Covariance between the neurons was determined from the firing rates of simultaneously recorded neurons over time. For each trial, the effective correlation[138] was calculated from the activities of the p neurons over time $\mathbf{X}$:

(45) $$\text{effective correlation} = 1 - \left\| \text{Corr}(X) \right\|^{1/p}$$

Effective correlations were then averaged across trials. Confidence intervals were determined from standard error between the sessions.

Code is in the directory https://github.com/dirkbeck/DM_space_model/tree/main/neuron_pair_analysis.

neuronSynchronyPlot.m generates the plot, which takes input files 'sessionDataOf....mat'.

These input mat files are generated by the 'covAndBinCtMatrixOfNeuronsFromSameSession.m' function which takes input files 'sameSession....mat'.

These input files are generated by 'extractNeuronsFromSameSession.m' function.

Changes to choice after adding cost to a reward offer.

In the analysis plotted in **Extended Data Fig. 3j,** we sought to define choice patterns before and after stress when a small cost was added to a reward offer. To do this, we analyzed choice data from the CBC task before and after stress (i.e. reward and a small cost) and in the BB task before after stress (i.e. only reward). We recorded the approach percentage across sessions and then averaged these between groups. Data from 17 rodents across 38 tasks is used for Control CBC, data from 13 rodents across 24 tasks for Stress CBC, data from 23 rodents across 114 tasks for Control BB, and data from 14 rodents across 116 tasks for Stress BB.

For data, see https://github.com/dirkbeck/DM_space_model/blob/main/disorder_hypotheses/experimental_data_analysis_choice_before_after_stress.xlsx.

Analyzed cortex-FSI connectivity.

Using data from the Corticostriosomal Circuit Stress Experiment database (published with Friedman et al. (2017)), firing rates and striatum-projecting prelimbic cortex neurons (i.e. sSPN-projecting cortex) and FSI were analyzed. For details regarding neuron classification, see Friedman et al. (2017). Spike data was converted to firing rates by separating data into 5-10 bins. A linear regression of the form $a \cdot x + b$ was fit through the firing rates. Examples are plotted in **Extended Data Figs. 3k,l** and averages of the square of Pearson correlation coefficient across neuron pairs and the a slope parameter from the regression fit are plotted in **Extended Data Figs. 3o,p**. 78 neuron pairs across 7 rodents were analyzed before stress and 37 neuron pairs were analyzed across 4 rodents after stress.

Code is in the directory https://github.com/dirkbeck/DM_space_model/tree/main/neuron_pair_analysis.

To produce **Extended Data Figs. 3k,l,** go to the directory https://github.com/dirkbeck/DM_space_model/tree/main/neuron_pair_analysis and run examplePairedNeuronPLot('twdb_control', 'PLSvsFSI', 54), examplePairedNeuronPLot('twdb_stress', 'PLSvsFSI', 32), respectively.

To produce **Extended Data Figs. 3o,p,** go to the directory https://github.com/dirkbeck/DM_space_model/tree/main/neuron_pair_analysis and run plotFitParamOfNeuronTriplets.m.

Modeled cortex-FSI connectivity after stress.

Connected cortical neurons and FSI show increased correlation and reduced slopes between their firing rates after stress (**Extended Data Figs. 3k,l,o,p**). To understand how this impacts SPNs, we modeled two changes to the cortex→FSI connectivity that might produce the experimental results (**Extended Data Figs. 3q,r**). A first factor, connection weight w_i, measured the strength of connection between the ith connected cortical neuron and the FSI, that is, how strongly FSI would respond to an increase in activity in the connected cortical neuron. A second factor, number of connections c, measured the number of connected cortical neurons to the FSI.

We solved for correlation between the neurons and slope using identities that link covariance, correlation, and slope. This process assumes that FSI receives input only from the cortex.

(46) $$\text{correlation}(\text{cortex 1}, \text{FSI}) = \frac{\text{cov}(\text{cortex 1}, \text{FSI})}{\sqrt{\text{var}(\text{cortex 1}) \cdot \text{var}(\text{FSI})}}$$

(47) $$\text{slope}(\text{cortex 1}, \text{FSI}) = \frac{\text{cov}(\text{cortex 1}, \text{FSI})}{\text{var}(\text{FSI})}$$

(48) $$\text{cov}(\text{cortex 1}, \text{FSI}) = w_1 \text{var}(\text{cortex 1}) + \sum_{i=2}^{c} w_i \cdot \text{cov}(\text{cortex 1}, \text{cortex } i)$$

(49) $$\text{var}(\text{FSI}) = \sum_{i=1}^{c} w_i^2 \text{var}(\text{cortex } i) + \sum_{i=1}^{c} \sum_{\substack{j=1 \\ j \neq i}}^{c} w_i w_j \text{cov}(\text{cortex } i, \text{cortex } j)$$

For simplicity, we assumed in our analysis that variance is equal across cortical neurons projecting to the FSI (i.e. $\text{var}(\text{cortex 1}) = \text{var}(\text{cortex 2}) = \ldots = \text{var}(\text{cortex } c)$) and that the covariance between those neurons is equal (i.e. $\text{cov}(\text{cortex 1}, \text{cortex } i) = \text{cov}(\text{cortex 1}, \text{cortex 2}) = \text{cov}(\text{cortex 1}, \text{cortex 3}) = \ldots = \text{cov}(\text{cortex 1}, \text{cortex } c)$). So, in the analysis plotted in **Extended Data Figs. 3q,r**, we substitute uniform variance and covariance to obtain:

(50) $$\text{correlation}(\text{cortex 1}, \text{FSI}) = \sqrt{\frac{\text{var}(\text{cortex 1}) + (c-1) \cdot \text{cov}(\text{cortex 1}, \text{cortex } i)}{c \cdot \text{var}(\text{cortex 1})}}$$

(51) $$\text{slope}(\text{cortex 1}, \text{FSI}) = \frac{1}{cw}$$

where:

- c is the number of cortex neurons connected to each FSI
- w is the weight of each cortex→FSI connection

Thus, changes in strength of connection may be more likely to affect slope, while changes in the quantity of connections, the covariance between cortical neurons, and the variance of cortical

neurons may also affect the cortex to FSI correlation. In stress, correlation and slope are both larger, suggesting reduction in quantity of cortex to FSI connections.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/disorder_hypotheses/ctx_to_FSI_synchrony_analysis.m.

Simulation illustrating the theoretical change.

We then conducted a simulation to examine the case where cortical data and cortex to FSI connectivities were less uniform. To do this, we constructed a network with random unit normal connection weights from cortex to FSI. Then we randomly lesioned all but several of the connections to reflect the sparsity of the brain. The connection from the first cortical neuron to the FSI was always preserved.

In each of ten simulations, a cortical input to sSPN was generated as a random unit normal 100x1 vector. Activity of the first cortical neuron and the FSI was recorded. These are plotted in **Extended Data Figs. 3m,n** along with a linear regression fit.

For code, see https://github.com/dirkbeck/DM_space_model/blob/main/disorder_hypotheses/ctx_to_FSI_synchrony_analysis.m.

Supplemental References

95. Lévesque, M. & Parent, A. The striatofugal fiber system in primates: A reevaluation of its organization based on single-axon tracing studies. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11888–11893 (2005).

96. Bevan, M. D., Bolam, J. P. & Crossman, A. R. Convergent Synaptic Input From the Neostriatum and the Subthalamus Onto Identified Nigrothalamic Neurons in the Rat. *Eur J of Neuroscience* **6**, 320–334 (1994).

97. Hajós, M. & Greenfield, S. A. Synaptic connections between pars compacta and pars reticulata neurones: electrophysiological evidence for functional modules within the substantia nigra. *Brain Research* **660**, 216–224 (1994).

98. Ragsdale, C. W. & Graybiel, A. M. Fibers from the basolateral nucleus of the amygdala selectively innervate striosomes in the caudate nucleus of the cat. *J of Comparative Neurology* **269**, 506–522 (1988).
99. Fujiyama, F., Unzai, T. & Karube, F. Thalamostriatal projections and striosome-matrix compartments. *Neurochem Int* **125**, 67–73 (2019).
100. Ragsdale, C. W. & Graybiel, A. M. Compartmental organization of the thalamostriatal connection in the cat. *J of Comparative Neurology* **311**, 134–167 (1991).
101. Unzai, T., Kuramoto, E., Kaneko, T. & Fujiyama, F. Quantitative Analyses of the Projection of Individual Neurons from the Midline Thalamic Nuclei to the Striosome and Matrix Compartments of the Rat Striatum. *Cereb Cortex* **27**, 1164–1181 (2017).
102. Martínez-Selva, J. M., Sánchez-Navarro, J. P., Bechara, A. & Román, F. [Brain mechanisms involved in decision-making]. *Rev Neurol* **42**, 411–418 (2006).
103. Zhou, J. *et al.* Evolving schema representations in orbitofrontal ensembles during learning. *Nature* **590**, 606–611 (2021).
104. Hamid, A. A. *et al.* Mesolimbic dopamine signals the value of work. *Nat. Neurosci.* **19**, 117–26 (2016).
105. Schultz, W. Dopamine reward prediction-error signalling: a two-component response. *Nat Rev Neurosci* **17**, 183–195 (2016).
106. Assous, M. & Tepper, J. M. Excitatory extrinsic afferents to striatal interneurons and interactions with striatal microcircuitry. *Eur J of Neuroscience* **49**, 593–603 (2019).
107. Ho, T. S. Y. & Lee, S. Term Structure Movements and Pricing Interest Rate Contingent Claims. *The Journal of Finance* **41**, 1011–1029 (1986).
108. Bergman, H. *et al.* Physiological aspects of information processing in the basal ganglia of normal and parkinsonian primates. *Trends in Neurosciences* **21**, 32–38 (1998).

109. Jhou, T. C. The rostromedial tegmental (RMTg) “brake” on dopamine and behavior: A decade of progress but also much unfinished work. *Neuropharmacology* **198**, 108763 (2021).
110. Baker, P. M., Rao, Y., Rivera, Z. M. G., Garcia, E. M. & Mizumori, S. J. Y. Selective Functional Interaction Between the Lateral Habenula and Hippocampus During Different Tests of Response Flexibility. *Front. Mol. Neurosci.* **12**, 245 (2019).
111. Nielson, H. C. & McIver, A. H. Cold stress and habenular lesion effects on rat behaviors. *Journal of Applied Physiology* **21**, 655–660 (1966).
112. Ilango, A. *et al.* Similar Roles of Substantia Nigra and Ventral Tegmental Dopamine Neurons in Reward and Aversion. *J. Neurosci.* **34**, 817–822 (2014).
113. Shabel, S. J., Proulx, C. D., Trias, A., Murphy, R. T. & Malinow, R. Input to the lateral habenula from the basal ganglia is excitatory, aversive, and suppressed by serotonin. *Neuron* **74**, 475–81 (2012).
114. Hyvärinen, A. Independent component analysis: recent advances. *Phil. Trans. R. Soc. A.* **371**, 20110534 (2013).
115. Gupta, A. *et al.* Neural Substrates of the Drift-Diffusion Model in Brain Disorders. *Front. Comput. Neurosci.* **15**, 678232 (2022).
116. Murray, J. D. Models for Interacting Populations. in *Mathematical Biology* (ed. Murray, J. D.) vol. 17 79–118 (Springer New York, New York, NY, 1993).
117. Wager, T. D. *et al.* A Bayesian Model of Category-Specific Emotional Brain Responses. *PLoS Comput Biol* **11**, e1004066 (2015).
118. Friedman, A., Keselman, M. D., Gibb, L. G. & Graybiel, A. M. A multistage mathematical approach to automated clustering of high-dimensional noisy data. *Proc Natl Acad Sci U S A* **112**, 4477–4482 (2015).
119. Abeles, M., Hayon, G. & Lehmann, D. Modeling Compositionality by Dynamic Binding of Synfire Chains. *J Comput Neurosci* **17**, 179–201 (2004).

120. Hayon, G., Abeles, M. & Lehmann, D. A Model for Representing the Dynamics of a System of Synfire Chains. *J Comput Neurosci* **18**, 41–53 (2005).
121. Berke, J. D. Functional properties of striatal fast-spiking interneurons. *Frontiers in systems neuroscience* **5**, 45 (2011).
122. McKeon, P. N., Bunce, G. W., Patton, M. H., Chen, R. & Mathur, B. N. Cortical control of striatal fast-spiking interneuron synchrony. *The Journal of Physiology* **600**, 2189–2202 (2022).
123. Sciamanna, G., Ponterio, G., Mandolesi, G., Bonsi, P. & Pisani, A. Optogenetic stimulation reveals distinct modulatory properties of thalamostriatal vs corticostriatal glutamatergic inputs to fast-spiking interneurons. *Sci Rep* **5**, 16742 (2015).
124. Peters, A. J., Fabre, J. M. J., Steinmetz, N. A., Harris, K. D. & Carandini, M. Striatal activity topographically reflects cortical activity. *Nature* **591**, 420–425 (2021).
125. Plenz, D. & Kitai, S. T. Up and Down States in Striatal Medium Spiny Neurons Simultaneously Recorded with Spontaneous Activity in Fast-Spiking Interneurons Studied in Cortex–Striatum–Substantia Nigra Organotypic Cultures. *J. Neurosci.* **18**, 266–283 (1998).
126. Humphries, M. D., Wood, R. & Gurney, K. Reconstructing the Three-Dimensional GABAergic Microcircuit of the Striatum. *PLoS Comput Biol* **6**, e1001011 (2010).
127. Liénard, J. & Girard, B. A biologically constrained model of the whole basal ganglia addressing the paradoxes of connections and selection. *J Comput Neurosci* **36**, 445–468 (2014).

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Corticostriosomal Circuit Stress Experiment	Friedman et al. (2017)	https://data.mendeley.com/datasets/z9jd8xhj84/1
A decision-space model explains context-specific decision-making	This paper	https://doi.org/10.7910/DVN/SMKW0I
Overview of the model	This paper	https://github.com/dirkbeck/DM_space_model/tree/main/model_overview
Tests of the model	This paper	https://github.com/dirkbeck/DM_space_model/tree/main/model_tests
Disorder hypotheses	This paper	https://github.com/dirkbeck/DM_space_model/tree/main/disorder_hypotheses
Instances 2 (sparse connectivity) and 3 (dynamics)	This paper	https://github.com/dirkbeck/DM_space_model/tree/main/dynamic_model_and_neural_net
Day to day differences and disorder comorbidity	This paper	https://github.com/dirkbeck/DM_space_model/tree/main/day_to_day_space_sampling
Circuit adjustment between trials	This paper	https://github.com/dirkbeck/DM_space_model/tree/main/circuit_trajectories
Analysis of the correlation between cortical neurons, FSIs, sSPNs, and mSPNs during decision-making	This paper	https://github.com/dirkbeck/DM_space_

		model/tree/main/neuron_pair_analysis
Analysis of the functional connectivity of sSPNs and mSPNs during decision-making	This paper	https://github.com/dirkbeck/DM_space_model/tree/main/Cross%20Correlation%20Pattern%20Counts
Software and algorithms		
MATLAB R2021a	Mathworks	https://www.mathworks.com/products/matlab.html