ARTICLE **Communicated by Eric T. Shea-Brown**

Neuromodulators Generate Multiple Context-Relevant Behaviors in Recurrent Neural Networks

Ben Tsuda
btsuda@mgh.harvard.edu
Stefan C. Pate
stefan.pate@northwestern.edu
Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037, USA

Kay M. Tye
tye@salk.edu
Systems Neuroscience Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037, USA

Hava T. Siegelmann
hava@umass.edu
Biologically Inspired Neural and Dynamical Systems Laboratory, School of Computer Science, University of Massachusetts Amherst, Amherst, MA, 01003

Terrence J. Sejnowski
terry@salk.edu
Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA 92037, USA; Institute for Neural Computation and Division of Biological Sciences, University of California San Diego, La Jolla, CA 92093, USA

Neuromodulators are critical controllers of neural states, with dysfunctions linked to various neuropsychiatric disorders. Although many biological aspects of neuromodulation have been studied, the computational principles underlying how neuromodulation of distributed neural populations controls brain states remain unclear. In contrast to external contextual inputs, neuromodulation can act as a single scalar signal that is broadcast to a vast population of neurons. We model the modulation of synaptic weight in a recurrent neural network model and show that neuromodulators can dramatically alter the function of a network, even when highly simplified. We find that under structural constraints like those in brains, this provides a fundamental mechanism that can increase the computational capability and flexibility of a neural network. Diffuse synaptic weight modulation enables storage

Ben Tsuda and Terrence Sejnowski are the corresponding authors.

Neural Computation **38, 292–327** (2026)
https://doi.org/10.1162/NECO.a.1489

of multiple memories using a common set of synapses that are able to generate diverse, even diametrically opposed, behaviors. Our findings help explain how neuromodulators unlock specific behaviors by creating task-specific hyperchannels in neural activity space and motivate more flexible, compact and capable machine learning architectures.

1 Introduction

Neuromodulators are a central mechanism in the biological control of neural states that manifest as mood, arousal, motivation, and other variable behavioral modes (Allen et al., 2019; Avery & Krichmar, 2017; Bacqué-Cazenave et al., 2020; Burgos-Robles et al., 2017; Inagaki et al., 2012; Katz & Edwards, 1999; Lee & Dan, 2012; McCormick et al., 2020). Other mechanisms, like exogenous contextual cues (e.g., a red light at a traffic signal), can also generate specific behaviors (Mante et al., 2013) by providing input signals that elicit specific activity patterns in a neural network. Although exogenous cue-based mechanisms can be advantageous in some circumstances (e.g., fast behavioral switching), they are impractical when massive scaling is needed. Broadcasting cue information across large brain regions would require many millions of neural connections to simply convey a constant cue signal to maintain conditioned activity (Dayan, 2012). Neuromodulation is an ideal system to control major changes of behavioral state, albeit at slower timescales, by modifying how neurons transduce information, including modulation of intrinsic ion channels and synaptic strengths. The impact of neuromodulation is not as specific as arrays of heterogeneous inputs, but we show here that it can flexibly select different computational outcomes in response to the same heterogeneous inputs.

A classic example of this trade-off is illustrated by sleep. Humans and many other species cycle through internal sleep stages that are controlled by a variety of neuromodulators. An exogenous cue-dependent system would require a "sleep cue" to be widely broadcast across the whole brain, constantly for hours, at great energetic expense. In contrast, neuromodulators that broadcast sparse signals widely can shift the entire brain into internal sleep states that persist for hours (Lee & Dan, 2012; McCormick et al., 2020).

The diffuse release of neuromodulators from a relatively small number of neurons projecting across huge regions of the brain can control a variety of "internal states" (Avery & Krichmar, 2017; Bargmann, 2012; Lee & Dan, 2012; O'Callaghan et al., 2020; Tarder-Stoll et al., 2020). In addition to sleep, neuromodulators are responsible for controlling mood and states of alertness (Gorfine et al., 2006; Lee & Dan, 2012) and regulating the consumption of food and water during awake states (Allen et al., 2019; Inagaki et al., 2012). Pioneering studies on the lobster and crab pyloric networks (Johnson et al., 1995; Marder, 2012; Marder & Eisen, 1984) and other systems (Jing et al., 2009; Katz et al., 1994) have revealed how

neuron-specific neuromodulation can precisely tailor central pattern generator rhythms. Yet it remains unknown how large-scale neuromodulation of vast distributed neural populations can control global network dynamics and dictate behavior as it does in large brains.

Fully understanding neuromodulation in brains is important for several reasons. First, most psychiatric disorders either stem from or are directly related to neuromodulator dysregulation, and many psychiatric drugs target neuromodulatory activity (Avery & Krichmar, 2017; Blokhin et al., 2020; Staudt et al., 2019). Second, many of the psychiatric drugs currently only partially or imprecisely target neuromodulatory processes and effect their results through mechanisms that remain unclear (Ashok et al., 2017; Bacqué-Cazenave et al., 2020; Hyman, 2013). Third, the effects of many psychiatric treatments are highly variable, with some patients responding strongly and others failing to respond to multiple drugs (Lauschke et al., 2019; Reynolds et al., 2013; Staudt et al., 2019; Voineskos et al., 2020). Fourth, neuromodulation acts via multiple mechanisms (as discussed below), allowing powerful circuit control but also making it difficult to fully understand their combinatorial actions (Avery & Krichmar, 2017; Marder & Thirumalai, 2002; Nadim & Bucher, 2012). Fifth, given the central role of neuromodulators in control of brains, a better understanding promises to make deep learning models based on brain architectures more flexible, more compact, and more efficient.

Neuromodulators affect several processes in brains including synaptic strengths, neural excitability, plasticity, and, indirectly, downstream circuit activity (Marder, 2012; Nadim & Bucher, 2012; Newman et al., 2012; Weele et al., 2018). Prior research has focused on different aspects of neuromodulation (Fellous & Linster, 1998), including Yu and Dayan (2005) who modeled the role of acetylcholine and norepinephrine in Bayesian probability estimations of uncertainties; Stroud et al. (2018) and Vecoven et al. (2020) who considered modulation of the neural activation function; Beaulieu et al. (2020) who formulated neuromodulation as a separate network that masks effector networks; Miconi et al. (2019) who used modulation of synaptic plasticity to train networks; Hasselmo et al. (1995) who developed a model incorporating experimental work on multifactor neuromodulator-specific circuit dynamics, particularly in hippocampal memory processes; and models of the locus coeruleus and norepinephrine (Aston-Jones & Cohen, 2005; Usher et al., 1999) that have helped elucidate the role of gain modulation in regulating decision making and balancing exploitation versus exploration in tasks. Each of these models has helped elucidate a component of neuromodulator function, helping form a more complete understanding of the various roles neuromodulators play.

We focus here on another critical aspect of neuromodulation in brains—synaptic weight modulation (Hasselmo, 2006; Johnson et al., 1995; Katz et al., 1994; Marder, 2012; Marder & Thirumalai, 2002; Nadim & Bucher, 2012)—a poorly understood and important control mechanism in brains.

We consider a simplified approximation in which neuromodulators act as nearly uniform synaptic weight amplifiers or dampeners—a single multiplicative scalar—within a local region of a neural network. We show how this form of neuromodulation establishes distinct memories using a common set of synapses, generating unique dynamic activity landscapes within a structurally conserved neural network. We demonstrate how neuromodulated circuits give rise to idiosyncratic, nonlinear dose-response properties that can differ depending on the mode of neuromodulation. Using a neuromodulation-mediated behavioral paradigm in *Drosophila*, we show how this form of neuromodulation naturally handles intermediate neural states and, as such, generalizes models of discrete internal state switching (Calhoun et al., 2019; Cermak et al., 2020) to continuous state transitions. Although many mechanisms may influence behavioral shifts, we show that a simple multiplicative factor applied to weights already acts as a powerful network control device, through which neuromodulators can increase the flexibility and complexity of computation in brains and suggest the potential for developing artificial neural networks that are more flexible, compact and capable.

2 Results

2.1 Neuromodulation Creates Multiple Weight Regimes within Shared Synaptic Connections. The effects of neuromodulators on synaptic weights present a mode of circuit control (Nadim & Bucher, 2012) that is poorly understood in brains—both how it is implemented at scale and the computational mechanisms by which it shifts coordinated activity to generate different behaviors. Several recent studies on cell type diversity have made clear that brains contain a complex array of neuromodulators that act with carefully coordinated spatiotemporal precision (Li et al., 2018; Patriarchi et al., 2018; Radnikow & Feldmeyer, 2018; Xie et al., 2021; Zhang et al., 2021). As a first step, we sought to assess whether a simplified form of neuromodulation—modeled as a uniform multiplicative factor acting on synaptic weights in a recurrent neural network (RNN)—could help us understand how neuromodulators control neural state. Although other modes of neural network control such as exogenous contextual cuing have been shown to successfully shift network behavior (Mante et al., 2013), uniform weight modulation represents a completely different biological and computational mechanism, which, given the complex, nonlinear, and often unpredictable nature of RNNs, requires explicit assessment. Furthermore, systems in which exogenous cues are required to drive neurons (Driscoll et al., 2024; Mante et al., 2013) are unsurprisingly more sensitive to external noise (see Extended Data Figure 2), while internally controlled neuromodulation-based systems are robust to such fluctuations in external inputs (Haddad & Marder, 2018) (see Extended Data Figures 1 and 2; although they are subject

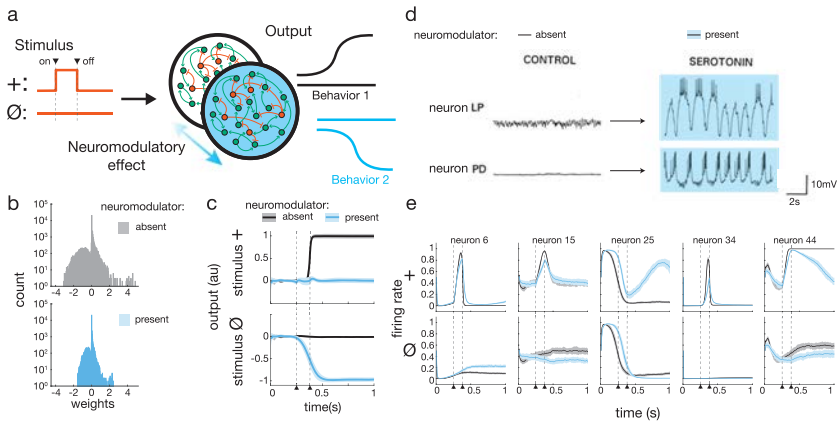

Figure 1: Neuromodulation weight scaling separates overlapping synaptic memory regimes. (A) Positive-negative output task. Given a stimulus (either + or ∅), in the absence of neuromodulatory effect, the recurrent neural network (RNN) should produce outputs from the Behavior 1 repertoire (input + → positive output; input ∅ → zero output) and in the presence of neuromodulator, from Behavior 2 (input + → zero output; input ∅ → negative output). (B) Neuromodulatory effect implemented in the model: all synaptic weights in the RNN are multiplied by a constant factor, here 0.5. (C) Mean output from 10 independently trained RNNs to + and ∅ input stimuli on the positive-negative output task with global neuromodulation factor 0.5. Shading represents standard deviation. (D) Individual neuromodulators elicit unpredictable transforms of intracellular voltage traces in crustacean stomatogastric ganglion (STG) neurons LP and PD. Blue-shaded traces indicate the presence of neuromodulator serotonin. (Reprinted and adapted from Neuron 76, Marder, Neuromodulation of Neuronal Circuits: Back to the Future, 1-11, 2012, with permission from Elsevier.) (E) Five example neurons' activity patterns from neuromodulated model RNN show complex nonlinear transformations analogous to crustacean STG activity changes under neuromodulation. Shading represents standard deviation. For A, C, and E, arrowheads with dashed lines indicate timing of onset and offset of stimulus pulse (for + stimulus).

to internal sources of noise: see Extended Data Figures 2c to 2f), supporting neural states like sleep as described above (McCormick et al., 2020).

We designed a simple experimental paradigm that we call the positive-negative output task (see Figure 1a and section 4) to assess whether given identical input stimuli, uniformly shifting all the recurrent weights in a RNN could elicit a unique trained behavior from the same network. In the positive-negative output task, one of two possible stimuli (+ or ∅) are input to the RNN. The RNN is trained to produce either a positive, zero, or negative output for each stimuli. For example, in Figure 1a, Behavior

1 in black shows output from an RNN trained to produce a positive output when given the positive (+) input signal and to produce zero output when given the zero (Ø) input signal. Under a different neuromodulation state, depicted in Figure 1a as the blue RNN and Behavior 2, the RNN produces different outputs to the same input stimuli (e.g., + stimulus elicits a zero output and Ø stimulus elicits a negative). This task simulates how a human or animal may behave differently to the same external cues while in a different internal state. For example, a thirsty mouse cued that only water is available (+ stimulus) may run to the water (positive output) and do nothing when no water cue is given (Ø stimulus leads to zero output), but a hungry mouse may do nothing when cued that only water is available (+ stimulus leads zero output) and forage for food when no water-only cue is given (Ø stimulus leads to negative output).

We found that neuromodulation simulated as broadcast synaptic weight scaling was able to generate the distinct behaviors for the task. For example, scaling all weights by a factor of 0.5 (see Figure 1b) could generate unique output behaviors (see Figure 1c). This simple mechanism demonstrates how neuromodulators operating in brains can effectively separate synaptic memory regimes within a fixed circuit and access them through uniform scaling of weights to unlock specific behaviors (see Extended Data Figure 4). We found this result held over a wide range of neuromodulatory factor magnitudes (see Extended Data Figures 5a to 5c). We refer to particular neural network weight configurations that generate particular behaviors as a "synaptic memory regime" since that configuration stores the memory to that specific input-to-behavior mapping. When the neuromodulator is applied, the same network transitions into a new synaptic memory regime, that is, a new configuration storing a distinct network memory specific to a different input-to-behavior mapping.

RNNs could successfully separate behaviors when recurrent weights only or recurrent, input, and output weights were all modulated at the same time, but not when only input or only output weights were modulated (see Extended Data Figures 5d to 5f). For the analyses presented below, we focus on modulation of only the recurrent weights.

Application of this form of neuromodulation led to unpredictable transformations of individual neuron activity patterns and global network activity (see Figure 1e and, Extended Data Figures 3a and 3b). This is reminiscent of the observed neuromodulator effects on individual neuron activity patterns in the stomatogastric ganglion of crabs (see Figure 1d) and other organisms (Johnson et al., 1995, 2011; Katz et al., 1994; Marder, 2012), where neural activities shifts have been shown to be unpredictable and display nonlinear transforms.

Given the stability of outputs toward the end of each trial in the positive-negative output task, we also investigated whether neuromodulated RNNs resulted in configurations that led to stable and reproducible fixed points as described in prior work (Driscoll et al., 2024). "Fixedpoint analysis" refers to a method of discovering points in activity state space where neural

dynamics are static when all inputs are held constant. This framework can prove useful in understanding principles of network dynamics that are reproducible between independently initialized and trained neural networks (e.g., attractor dynamics, decision boundaries, unstable points as in Driscoll et al., 2024). These dynamics can then be studied in-depth as inputs, network parameters, or training parameters are varied. We applied fixed-point analysis (Golub & Sussillo, 2018) to neuromodulated RNNs trained on the positive-negative output task (see section 4 for details of analysis). Interestingly, we found that neuromodulated RNNs did not consistently yield fixed points (see Extended Data Figure 6). Fixed-point analysis is dependent on specific network design, hyperparameters, and task and in our case did not yield interpretable results. This inability to identify consistent fixed points between replicate RNNs precluded further application of this analysis technique.

2.1.1 Targeted Neuromodulation Can Toggle Behaviors across Multiple Global Network States. In brains, neuromodulators are released in specific regions—some tightly localized, others broadcast widely—to influence local and global neural output. We found that RNNs with neuromodulated subpopulations of sizes across a broad range (100%–10% of the whole population) consistently supported the opposing behaviors of the task (see Figures 2a to 2c). Just as some neuromodulators affect neurons in a cell-type specific manner, for example, selectively influencing activity of excitatory or inhibitory neurons with corresponding receptors (Wester & McBain, 2014), we found targeting of neuromodulators in this manner also was able to generate multiple behaviors (see Figures 2d and 2e).

We next tested whether several unique behaviors could be learned and unlocked from a single network through targeted neuromodulation. To assess this, we created an extended version of the positive-negative output task in which the same two input stimuli (+ and ∅) could elicit nine unique combinations of corresponding output behaviors from an RNN (see section 4 for details). We found that neuromodulation of distinct subpopulations or with distinct neuromodulation levels could support the multiple output behaviors. In a single 200-unit neural network, we were able to generate nine distinct behaviors by varying the location of where the neuromodulator effect was broadcast (see Figures 2f and Extended Data Figures 7 and 8).

2.1.2 Distinct Global Network Activity Hyperchannels Emerge from Neuromodulation and Exhibit Nonlinear Transition Dynamics. To understand how this form of neuromodulation leads to network behavior shifts, we analyzed the coordinated activity of all neurons in the RNN in the absence and presence of neuromodulators. At the individual neuron level, neuromodulation shifted the net difference of excitatory and inhibitory inputs, which in turn altered the recurrent propagation of activity over time and resultant internal network dynamics (see Extended Data Figures

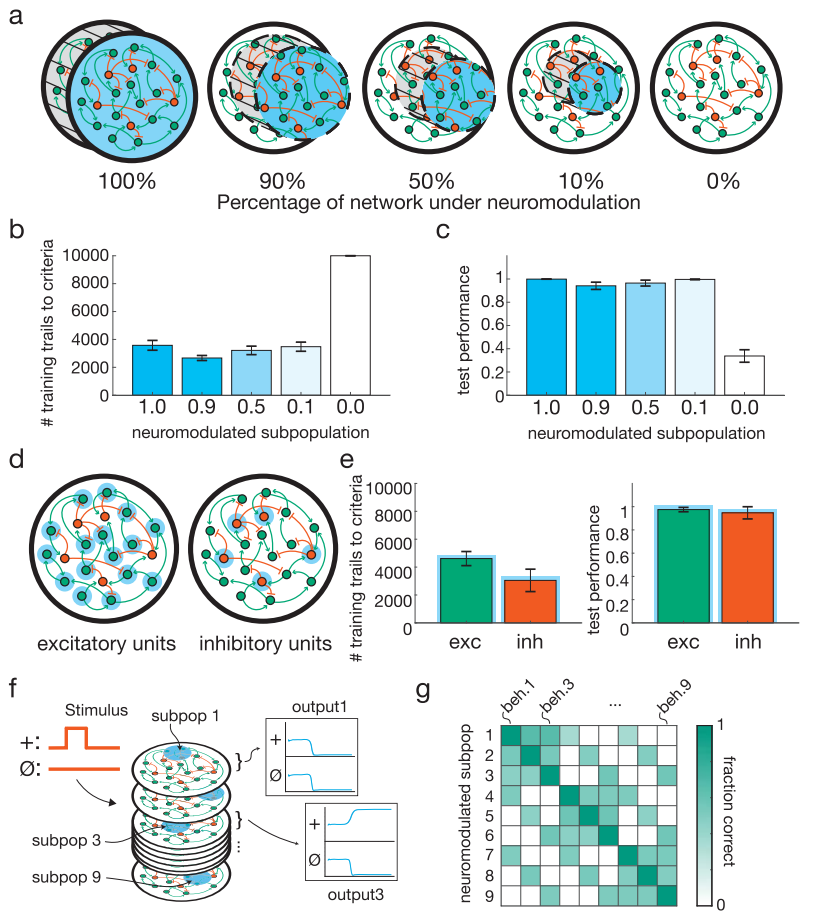

Figure 2: Targeted neuromodulation flexibly supports multiple behaviors. (A) A range of different-sized neural subpopulations (0%–100% neural units) embedded within an RNN as neuromodulated (factor (f_{nm}) = 0.5). 100% was positive control demonstrated in Figure 1; 0% was negative control. (B–C) RNNs with embedded neuromodulated subpopulations across the size spectrum could support the behaviors of the positive-negative output task. (B) Number of training trials required to reach stop criteria, indicating successfully trained networks (see section 4). (C) Test performance of trained RNNs (1 on y-axis is 100% correct; see section 4). (D) Neuromodulation of exclusively excitatory or inhibitory neurons (blue annuli). (E) Excitatory or inhibitory neuromodulation supported learning of the positive-negative output task. (F) Nine-behavior extended positive-negative output task with unique neuromodulated subpopulations (subpops) and example corresponding outputs. (G) RNNs successfully learned the task from panel F with nine targeted subpops (each 10% of the RNN, nonoverlapping; $f_{nm} = 2.5$). Application of the

9 to 11). At the whole population level, neural activity trajectories for the same stimulus with and without neuromodulator followed nonoverlapping, stereotyped paths, or hyperchannels (Goudar and Buonomano, 2018; Low et al., 2018; Nieh et al., 2021), in activity space (see Figure 3a). Activity hyperchannels, also called "hypertubes" in the literature (Goudar & Buonomano, 2018), refer to bundles of stereotyped trajectories in activity space (activity patterns) that define the constrained subspace of all possible network activities that is occupied when the network is in a given state (e.g., the neuromodulation level) and presented with a given stimulus (e.g., positive stimulus), as defined in Goudar and Buonomano (2018). Through amplification of synaptic weights, neuromodulation effectively resets all relative synaptic weights, altering activity flow patterns through the RNN (see Extended Data Figures 1c to 1d and 12a).

The distinct hyperchannels in activity space derive from a common underlying neural network. This suggests that there must be a transition between the hyperchannels that is accessible through intermediate amounts of neuromodulation. To characterize these transitional states, we applied intermediate levels of neuromodulation to the RNN after training, that mapped a smooth transition from trajectories of the nonneuromodulated hyperchannel to those of the fully neuromodulated hyperchannel (see Figure 3b). We found that intermediate neuromodulation generated intermediate outputs from the network (see Figure 3c). In this way, neuromodulated neural trajectories provide a means to dynamically respond to intermediate, even unexperienced, neural states (e.g., hunger levels), just as in other work, RNN neural trajectories have been shown to naturally address temporally varying sensory-motor patterns (Goudar & Buonomano, 2018).

We then characterized how the RNNs' output behaviors transitioned over the spectrum of applied intermediate neuromodulation levels. To compare different RNNs' output behaviors, we took the output of each RNN at the midpoint of each trial for each level of neuromodulation, which captured each RNN's output behavior response after any given stimulus presentation. We found that increasing neuromodulator levels led to nonlinear progression from nonneuromodulated behavior to fully neuromodulated behavior (see Figure 3d; nonneuromodulated behavior in this simulation was output of $+1$; neuromodulated behavior was output of 0; behavior progression was best fit by exponential or sigmoid function).

---

neuromodulator to any subpop unlocked a specific behavior set (beh.) from the nine-behavior repertoire (fraction of trials correct is approximately 1 on diagonal; see section 4). Off-diagonal fraction correct due to partial output overlap between behavioral sets (e.g., $\emptyset$ output as shown for output behaviors 1 and 3 shown in panel F is the same); for perfect performance, off-diagonals should be 0.5 or 0 depending on whether there is output overlap.

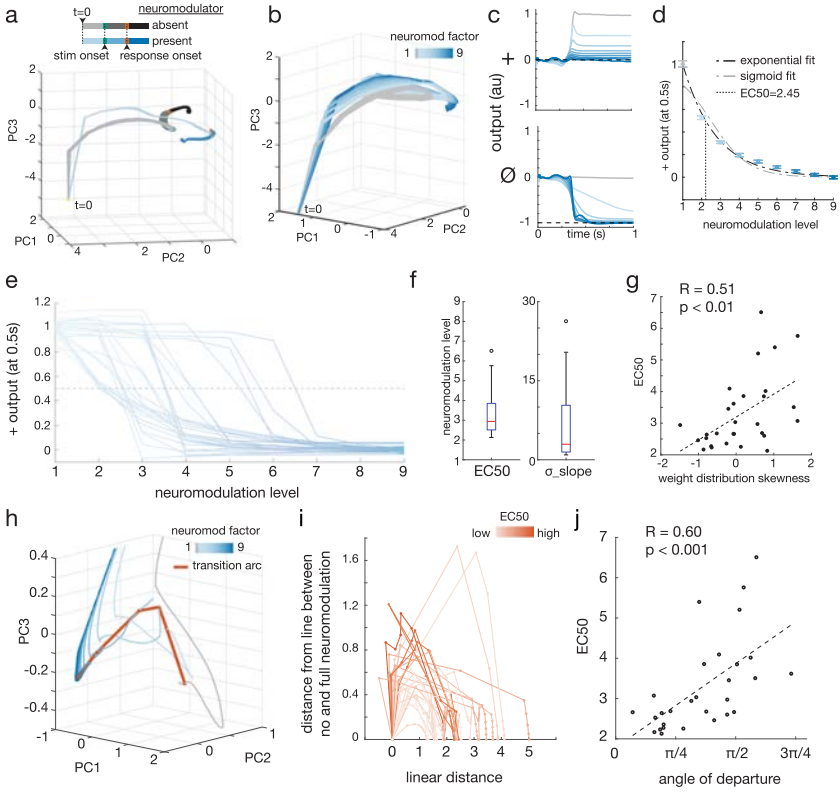

Figure 3: Neuromodulation separates activity hyperchannels with idiosyncratic nonlinear transition dynamics. (A) Global network activity dynamics after PCA-based dimensionality reduction for positive stimulus. Activity trajectories follow stereotyped, nonoverlapping paths or hyperchannels in activity space. (B) For an RNN trained with a neuromodulatory factor of 9, intermediate levels of neuromodulation lead to partial transitions (traces in shades of blue) toward full neuromodulation activity hyperchannel (traces in darkest blue). (C) Partial neuromodulation maps to intermediate output behaviors. (D) Transition from none to full neuromodulation behavior is nonlinear (best-fit exponential and sigmoid shown) and defines network sensitivity, measured as EC50. (E) Output transition for 29 networks independently trained on the positive-negative output task and tested across intermediate levels of neuromodulation. (F) Twenty-nine independently trained RNNs exhibit large variability in transition dynamics, with EC50 ranging from 2.1 to 6.5 and sigmoid slope (σ_slope) from 0.9 to 26.3. Network EC50 is positively and significantly correlated with skewness of global weight distribution. (H) Zoomed image demarcating path at a given time point over a network's neuromodulation transition manifold (orange; transition arc) that connects hyperchannels corresponding to different neuromodulation levels in activity space (gray–blue trajectories). Blue shading

We next assessed the degree of variability of this neuromodulator-dependent nonlinear behavioral progression. It could be that similar neuromodulation of different networks results in similar transitions in behavior, representing a conserved property of neuromodulation transitions. Alternatively, individual differences in network connectivity could cause high variability in behavioral transition dynamics under neuromodulation. To evaluate this, we independently trained several RNNs generated with identical parameter constraints ($n = 29$; see section 4 for parameter constraints used). All RNNs exhibited nonlinear transition dynamics best fit by an exponential or sigmoid function (see Figures 3e and Extended Data Figure 12b), but networks' sensitivities to neuromodulator and rates of transition varied drastically. To quantify this variability, we defined a "half maximal effective concentration" (EC50) as the amount of neuromodulator required to generate a half maximal output, analogous to the metric used to assess drug efficacy in pharmacology (see section 4). While all networks could produce intermediate outputs, the EC50 of individual networks trained with a full neuromodulator factor of 9 ranged from 2.1 to 6.5 (3.1× range; see Figure 3f, left) and rate of transitions (steepness of the transition dynamics sigmoid) varied widely as well from 0.9 to 26.3 (see Figure 3f, right). This result reveals a new, previously uncharacterized form of "circuit-based sensitivity" (Alonso & Marder, 2020; Dethier et al., 2015; Gutierrez & Marder, 2013; Jang et al., 2012) that occurs even when all parameters of the circuit are selected in an identical manner and networks exhibit identical end behaviors. This variability, which is derived purely from the stochastic nature of weight initialization (the random starting configuration) and specific training experience (the specific sequence of instances experienced) may contribute to the wide individual variability of neuropsychiatric drug sensitivities observed clinically (Furukawa et al., 2019; Staudt et al., 2019; Voineskos et al., 2020).

---

same as in panel B. At a given time point (here $t = 100$), intermediate levels of neuromodulation (represented by individual points) trace an arc between no neuromodulation (left-most point on each curve) and full neuromodulation (right-most point on each curve) states in phase space. Purely linear interpolation would lie along the x-axis at "distance from line (y-axis)" $= 0$. Each curve represents an individual network's transition arc across neuromodulation levels. Arcs are shaded by relative EC50. Arcs represent cross-section of full transition manifold connecting hyperchannels across all time points at intermediate neuromodulation; each point on the arc is an average of a cross section of a single hyperchannel. (J) The angle of departure (angle formed between direct path and first neuromodulation-level hyperchannel) exhibits a strong positive correlation with network EC50. Transition in direction more orthogonal or away from full neuromodulation state results in lower sensitivity, that is, higher EC50. For all PCA plots, the top three PCs accounted for 80% to 92% of activity variance.

We next sought to characterize what properties of the networks contribute to the variability in sensitivity to neuromodulator. Since the neuromodulation effect we simulated acted on synaptic weights, we first analyzed individual network weight distributions. We found that the skewness of networks' weight distributions exhibited a positive correlation with EC50s ($R = 0.51$, $p < 0.01$; see Figure 3g), suggesting networks with more positively skewed weights (longer tail of strong excitatory weights) were less sensitive to neuromodulators.

To further understand the source of network sensitivity variability, we characterized the shape of the networks' activity transition curves across neuromodulation levels (see Figure 3h). At a given trial time point, purely linear interpolation yielded linear sensitivity relationships with invariant EC50 (see Extended Data Figure 13). In contrast, progressive neuromodulation defined an arc (see Figures 3h and 3i). Taken across all time points, this arc extended to form a curved transition manifold connecting each activity hyperchannel defined by a specific intermediate neuromodulation level. "Manifold" here refers to a constrained region of activity state space that the network's activity occupies and transitions across depending on activity initialization parameters, stimuli, and neuromodulation level. The geometry of each network's transition arc (measured as the angle of departure; see section 4) was strongly correlated to network sensitivity (see Figure 3j). This suggests that while individual networks achieve identical performance on the trained task (at no and full neuromodulation), the geometry of their population activities at intermediate neuromodulation levels is unique, leading to emergent sensitivity profiles. In particular, characterization of networks' transition arcs revealed that transition dynamics are nonlinear due to curvature of the activity space transition, and that the specific shape of a network's transition arc is tightly linked to its sensitivity to neuromodulator.

Excessive neuromodulation can also occur either pathologically or pharmacologically. To model this, we applied neuromodulation at levels higher than those used during training and found neural dynamics could sometimes diverge from trained activity hyperchannels into an adjacent region of activity space, translating into inappropriate output behavior (see Extended Data Figures 12c to 12e).

2.1.3 Neuromodulated RNN Replicates Dopamine-Mediated Starvation-Dependent Sugar Sensitivity in Drosophila. Given our finding that neuromodulation provides a natural means of handling intermediate, unexperienced neural states, we next sought to evaluate our model's ability to recapitulate the behavioral effects of neuromodulation observed in vivo. The neuromodulator dopamine controls the sugar sensitivity behavior of *Drosophila* (Inagaki et al., 2012), as measured by proboscis extension reflex (PER) probability, which increased with both duration of starvation (fed, 1-day-starved; 2-day-starved) and concentration of L-dopa administered

in their diet (0, 3, 5 mg/ml) (see Figure 4a). PER is a behavior of *Drosophila* in which they extend their proboscis in response to a desirable food stimulus.

To assess if our neuromodulation model could reproduce these results, we trained RNNs with neuromodulated subpopulations (20% subpopulation) to reproduce the fed and 2-day starved sugar sensitivity curves of flies (no neuromodulation ($f_{nm} = 1$) for fed; $f_{nm} = 5$ for 2 day starved). We then tested the RNNs' behaviors at an intermediate, never-before-experienced neuromodulator level ($f_{nm} = 3$). The RNNs produced a shifted sensitivity curve very similar to that exhibited by 1-day starved flies, and the flies fed an intermediate L-dopa concentration of 3 mg/ml (see Figure 4b), as measured by the sugar sensitivity curve shift and mean acceptance threshold (MAT), which is the sugar concentration at which 50% of flies or networks exhibit the PER (see section 4).

The behavior of the RNNs reliably mimicked the intermediate behaviors of flies in vivo because intermediate neuromodulation caused a continuous shift in the RNN's activity hyperchannel between "fed" and "2-day starved/5 mg/ml L-dopa" hyperchannels (see Extended Data Figures 14a to 14c). Furthermore, our model reveals that transition manifolds (defined by the hyperchannels connecting intermediate neuromodulation levels) are unique to networks, predicting wide-ranging natural variability in fly starvation-based sugar sensitivity profiles (see Extended Data Figure 14d). Neuromodulation in our model leads to natural handling of never-before-experienced neural states by creating a network configuration such that the neuromodulatory transition manifold has a geometry that leads to intermediate outputs. Although the dopamine-dependent control of sugar sensitivity in *Drosophila* is likely regulated by other effects as well, our model shows that single scalar synaptic scaling is already sufficient to fully account for the starvation-dependent sugar sensitivity shifts observed.

2.1.4 Electrical Modulation Shifts Neural Dynamics through an Independent Circuit Effect. Other endogenous and exogenous, influences can alter neural circuit dynamics through mechanisms that may be shared or independent, and understanding relationships between such interventions is vital for safe and effective treatment. We used our model to compare whether exogenously delivered electrical modulation of a neuromodulated circuit—analogous to use of optogenetics experimentally (Deisseroth, 2015) and deep-brain stimulation (DBS), transcranial magnetic stimulation (TMS), or transcranial direct current stimulation (tDCS) clinically (Gouveia et al., 2019; Krauss et al., 2021; Risio et al., 2020; Thair et al., 2017)—alters network activity in an analogous manner to chemical neuromodulation or operates through an independent effect (see Figure 5a).

We simulated electrical modulation as added excitatory or inhibitory current (see section 4) to RNNs that had been trained with neuromodu-

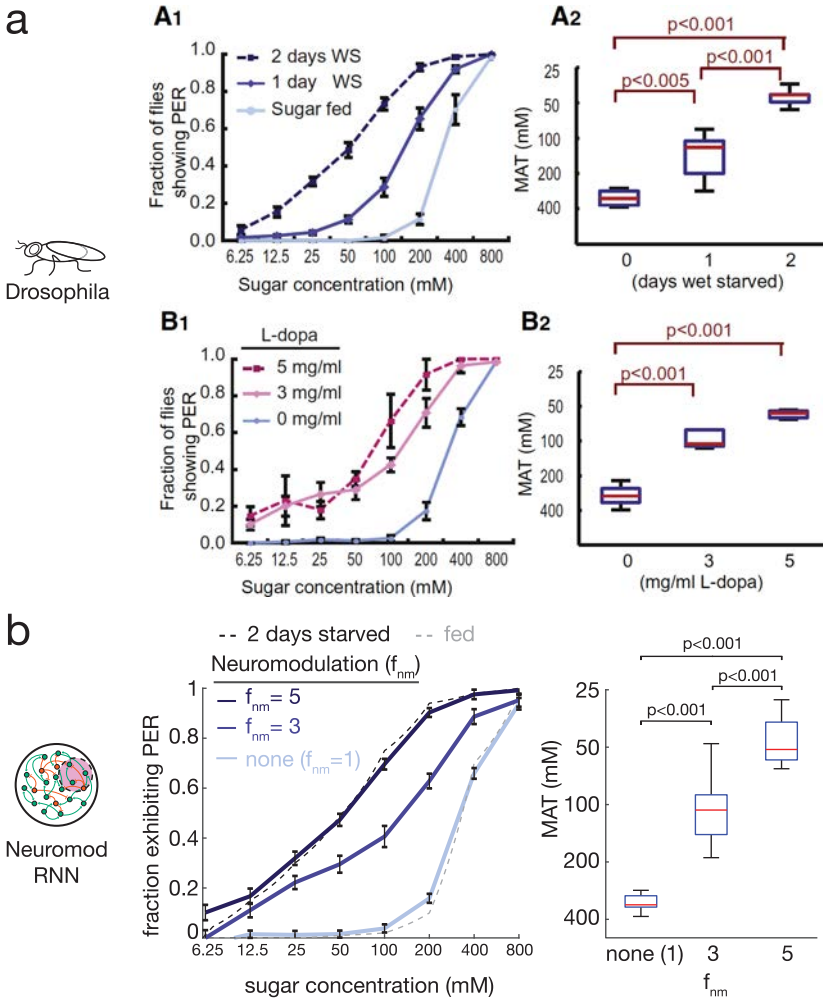

Figure 4: Neuromodulated RNNs reproduce *Drosophila* sugar sensitivity behaviors. (A) *Drosophila* sugar sensitivity behaviors from Inagaki et al. (2012) measured as proboscis extension reflex (PER) behavior versus sugar concentration and mean acceptance threshold (MAT; the sugar concentration that elicits 0.5 PER fraction). (Reprinted from Cell 148, Inagaki et al., Visualizing Neuromodulation In Vivo: TANGO-Mapping of Dopamine Signaling Reveals Appetite Control of Sugar Sensing, 583–595, 2012, with permission from Elsevier.) (B) Neuromodulated RNNs trained on extremes of *Drosophila* sugar sensitivity (no neuromodulation factor ($f_{nm} = 1$) for fed and $f_{nm} = 5$ for two-days starved) exhibit similar intermediate ($f_{nm} = 3$; untrained) and extreme (no neuromodulation ($f_{nm} = 1$) and $f_{nm} = 5$; trained) behaviors ($n = 10$; error bars are SEM; same statistical test as in Inagaki et al. (2012) for boxplots; see section 4).

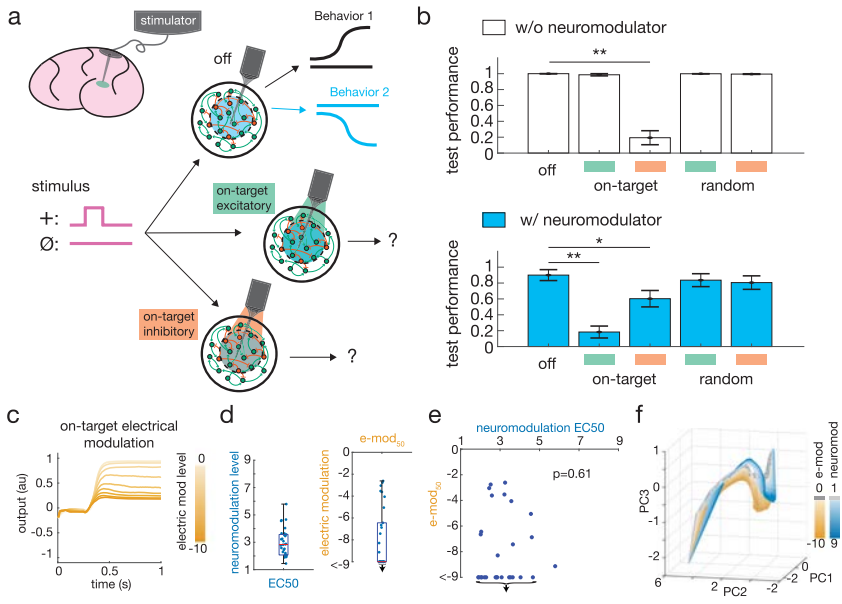

Figure 5: Targeted electrical modulation shifts network dynamics through independent circuit effect. (A) Schematic of DBS and analogous electrical modulation (e-mod) of a neuromodulated RNN on the positive-negative output task. "Off-target" refers to stimulation—either excitatory (green) or inhibitory (orange)—applied to a random subpopulation of neurons in the network of equal size to the size of the neuromodulated subpopulation; "on-target" refers to stimulation applied specifically to the neuromodulated subpopulation of neurons. (B) Test performance was significantly impaired in the absence of neuromodulation (white bars) when inhibitory on-target electrical modulation (e-mod) was given ($p = 3.91\text{e-}11$) and with neuromodulation (blue bars) when excitatory and inhibitory on-target e-mod was given ($p = 2.20\text{e-}08$ and $p = 2.18\text{e-}02$, respectively). $^{*}p < 0.05$, $^{**}p < 0.01$, Student's t-test. Green indicates applied excitatory e-mod; orange indicates inhibitory e-mod. Performance decreased due to behavioral shifting directionally toward opposing neuromodulation condition (see Extended Data Figures 15a and 15b). Error bars are SEM over 10 independent networks (see section 4). (C) Example RNN output in absence of neuromodulation to + stimulus with increasing inhibitory e-mod. (D) For 30 RNNs: Left: neuromodulation EC50. Right: electrical e-mod$_{50}$. Down arrows in (D–E) indicate RNNs that did not achieve e-mod$_{50}$ at maximum stimulation. (E) No significant correlation between networks' EC50 and e-mod$_{50}$. (F) Neuromodulation and electrical modulation push activity trajectories along different transition manifolds enabling independent transition dynamics.

lated subpopulation (50% of neurons) on the positive-negative output task. We found that while electrical current given to random subpopulations ("off-target" stimulation) did not affect the RNNs' performances, current delivered to the neuromodulated subpopulation ("on-target" stimulation) could significantly affect network output (see Figure 5b), shifting behaviors directionally toward the opposing neuromodulation condition behavioral set (see Extended Data Figures 15a and 15b).

In each RNN, increasing the level of targeted electrical input (e.g., inhibitory modulation in the absence of neuromodulator) led to a graded transition in behavior similar to the transition observed with graded administration of neuromodulator (see Figure 5c and section 4). To compare these conditions, we measured the amount of electrical input that led to output of 50% of the fully neuromodulated condition (e-mod$_{50}$) analogous to EC50 for neuromodulator levels. Interestingly, RNNs also exhibited idiosyncratic circuit-based sensitivity to electrical modulation, with e-mod$_{50}$ under inhibitory modulation ranging from -2.6 input current to not achieving e-mod$_{50}$ by the maximum modulation we administered (-9 input current; $> 3.5\times$ range) (see Figure 5d, right).

We then assessed whether networks' electrical and chemical sensitivities were related and found no significant correlation (see Figure 5e). Since some RNNs' outputs did not reach e-mod$_{50}$, saturating before the maximum electrical input given, further evidence of a different mechanism at play, we also measured each RNN's output at maximum electrical input (input = –9) and similarly found no significant correlation to EC50 (see Extended Data Figure 15c). The lack of correlation suggests that networks insensitive to chemical modulation may still be highly sensitive to electrical modulation and vice versa. Consistent with this, we found that electrical modulation progressively shifted population dynamics along a manifold transition distinct to neuromodulation, enabling different rates of transition (see Figure 5f). This may help explain why some patients who fail pharmacological treatments sometimes respond dramatically to electrical neuromodulation interventions like DBS or TMS. In addition to their localized effects to specific brain regions, DBS and TMS utilize a different mechanism to chemical modulation, leveraging an independent circuit-sensitivity to achieve therapeutic efficacy.

3 Discussion

Neuromodulation in brains drives unique neural function in health and disease. Using an RNN model, we showed how neuromodulators can act through simple scaling of synaptic weights to generate unique behavioral modes from a single RNN. Our model sheds light on how neuromodulation works biologically, specifically neuromodulator effects on synaptic weights (Hasselmo, 2006; Johnson et al., 1995; Katz et al., 1994; Marder, 2012; Marder & Thirumalai, 2002; Nadim & Bucher, 2012). Using a series of tractable tasks,

we showed how this unique and previously uncharacterized control mechanism in RNNs is able to rapidly reconfigure a network with only a single scalar input.

We make two specific predictions. First, we posit that neural networks exhibit idiosyncratic activity transition profiles resulting in varying sensitivities to neuromodulation. Networks' sensitivities depend on the geometry of their transition profiles and weight distributions, specifically the shape of the transition manifold defined in activity space. Second we predict that chemical neuromodulation acts on networks via a separate transition manifold to that of electrical stimulation, resulting in uncorrelated sensitivity profiles between these different modes of neuromodulation.

These predictions can be directly tested. Although there is already indirect evidence that neural networks in the brain have idiosyncratic sensitivities to neuromodulation via variable response to treatments like SSRIs, direct evidence could be gained with experiments applying varying amounts of neuromodulator (e.g., concentrations of serotonin) to a specific circuit, measuring network output, and repeating this over several circuits (in vitro) or organisms (in vivo) to compare their response profiles. Neural recordings of the circuits during these experiments would enable assessment of transition profiles in neural activity space to test our hypothesis about transition manifold geometries. At a behavioral level, our simulations studying *Drosophila* behavior could be tested experimentally by measuring the behavioral transitions of individual flies and measuring the variability compared to our model's prediction. Our second prediction could be tested with experiments applying both chemical neuromodulation (e.g., application of serotonin to a circuit) as well as direct electrical stimulation (via electrode or optogenetics). By varying the degree of each mode of neuromodulation, output response profiles can be measured and compared to characterize if each mode of neuromodulation operates along an independent, uncorrelated transition profile as predicted by our simulations. Here again, simultaneous neural recording would enable activity space characterization and comparison to our model's transition manifold predictions.

Our model provides insights into neuromodulation that are related to other recently elucidated principles of neural computation. We showed that neuromodulation leads to separation of distinct activity hyperchannels, similar to those observed by Goudar and Buonomano (2018) and Nieh et al. (2021), with neuromodulation effectively disentangling neural trajectories by separating them in phase space analogous to the work of Russo et al. (2018) in motor cortex. Just as neural trajectories provide transformations in phase space that naturally handle temporal variation of sensorymotor patterns (Goudar & Buonomano, 2018), neuromodulation leads to transformations in phase space that elucidate a biological mechanism for handling intermediate and continuously transitioning neural states, even if never experienced before. We demonstrated the biological use of this property through replication of Inagaki et al.'s (2012) findings in *Drosophila*. In this way, the level of neuromodulation acted as a controller on the amount

of disentangling of neural trajectories, using internal neural state (amount of weight modulation) to control output behavior. Such a system is robust to external noise (see Extended Data Figure 2), since far apart neural states generate trajectories that are widely separated in phase space.

Additionally, our model of synaptic weight modulation shares some similarities to previous models, particularly to the neural excitability models (Stroud et al., 2018; Vecoven et al., 2020). Both lead to increased flexibility and versatility, yet they operate through independent mechanisms biologically (Nadim & Bucher, 2012) and computationally (see appendix A in the Extended Data).

A recent work by Driscoll et al. (2024) presented an exogenously context-cued RNN model in which they smoothly varied specific stimulus inputs and characterized network activity and output, somewhat analogous to our analysis of smoothly varying intermediate neuromodulation levels. They performed detailed fixed-point analysis to characterize their RNNs and found two interesting and relevant results. First, when they varied the response-stimulus parameter and tracked fixed-point variation, they found a smoothly rotating fixed-point geometry. This indicates that exogenously cued networks can also follow smooth transition dynamics, though notable in their analysis, these stimuli had been trained on and thus were not novel to the network. In a second analysis, they showed that varying the context-stimulus led to movement of stable and unstable fixed points such that a decision boundary was defined, in this case for context stimuli not experienced during training. This result highlights the importance of the task along with other network parameters (e.g., "g" gain, training procedure, initialization) in defining the network dynamics and occurrence or lack of bifurcations; that is, particular task and network conformations may lead to smooth or bifurcating dynamics, as we also saw when neuromodulator levels outside the training range were applied (rather than intermediate levels; see Extended Data Figures 12c to 12e). Future detailed comparison of multiplicative synaptic-weight-scaling neuromodulated RNNs, gain-modulated RNNs, exogenously context-cue driven RNNs, and RNNs incorporating all of these mechanisms applied to various task types will be informative.

Through our analysis, we discovered a type of circuit-based sensitivity previously unreported to our knowledge. Previous studies have characterized network parameters (e.g., neuron types (Jang et al., 2012), maximal gain conductances (Alonso & Marder, 2020; Dethier et al., 2015; Goldman et al., 2001), and specific rectifying currents (Gutierrez & Marder, 2013)) that can alter sensitivity of a network when varied and demonstrate that individual networks can behave differently. Here, we identify a form of circuit-based sensitivity that arises when networks are created under exactly the same conditions (no network parameter is systematically varied between them) and generate exactly the same end behaviors. The variability in sensitivity emerges only in their transition dynamics and results purely from the stochastic nature of network weight initialization and unique individual experiences (random sequences of the exact same training data). This

type of circuit-based sensitivity may contribute (along with other forms) to the high variability in drug and other therapeutic response observed clinically (Furukawa et al., 2019; Staudt et al., 2019; Voineskos et al., 2020), alongside more standard explanations like enzyme variant-dependent drug metabolism and clearance rates (Lauschke et al., 2019; Ozomaro et al., 2013; Reynolds et al., 2013; van Westrhenen et al., 2020).

This emergent sensitivity property of neuromodulated networks is related to the well-known "many solutions" phenomenon where different weight configurations can produce identical output (Maheswaranathan et al., 2019; Mehrer et al., 2020; Prinz et al., 2004). Our model demonstrates a new facet of circuit variability derived from the transition dynamics in RNNs under varying neuromodulation. We reveal that unique geometric configurations of phase space underlie emergent network sensitivity profiles. An interesting future research question is whether the sensitivity described in this study is related to other forms of circuit sensitivity previously described, for example, the variability observed in 2 to 16 neuron models comprising Hodgkin-Huxley-type neurons initialized with different maximal gain conductances. Are particular weight profiles in the larger networks analogous to a range of maximal conductance parameters as was investigated in these smaller network models (Dethier et al., 2015), perhaps also relating to "sensitive" and "insensitive" directions in both parameter space (Goldman et al., 2001) and activity space?

Future studies aimed at identifying circuit parameters that control neural activity transition dynamic geometry will be critical for understanding and use in therapeutic optimization. New insights into network sensitivity will be gained by developing useful metrics to characterize how a given network's activity space is arranged (the shape of the transition surface) and understanding the relationship between networks' synaptic weight profiles and their activity space geometries. Our study begins to contribute to this effort by identifying the activity transition surface geometry as a critical factor in predicting network sensitivity. The idiosyncratic nature of circuit-based sensitivity aligns with current efforts in precision medicine calling for the need to consider each patient as an idiosyncratic individual. Here we provide computational evidence for this claim and its particular importance in neuropsychiatric treatment (Provenza et al., 2019). Fully understanding the relationship between chemical and electrical modulation and sensitivity is also crucial. Although our simplified model suggests how the modes of modulation influence dynamics (see appendix B in the Extended Data), further analytical and experimental investigation into their relationship as network dynamics evolve over time could provide deeper insights.

Our formulation of the neuromodulatory effect on synaptic weights is a simplification of the true biological mechanism. Elaborating our model to support differential weight modulation (e.g., via multiple neuromodulators and neuromodulator receptor subtypes on specific cell types) (Bacqué-Cazenave et al., 2020; Disney, 2021; Harris-Warrick & Johnson, 2010),

neuromodulator multiplexing (Li et al., 2018; Wester & McBain, 2014), and metamodulation (Katz & Edwards, 1999; Nadim & Bucher, 2012; Ribeiro & Sebastiao, 2010) will likely lead to even more sophisticated network behavior. Our model also uses arbitrary neuromodulation levels, whereas brains likely use specific levels for optimal functionality. Future investigation into which levels are optimal and methods of learning these will be important.

Additionally, the tasks we used to investigate how this aspect of neuromodulation works were relatively abstract and simple: scalar range to scalar range mapping in the *Drosophila* task and dynamic time-varying signal to time-varying output in the positive-negative output task. This allowed us to tractably demonstrate and probe properties of neuromodulation and how they affect network computation and behavior. Future investigation into whether this simple mechanism alone is sufficient to support libraries of much more complex behaviors within a network, or context-dependent behaviors in other learning paradigms such as reinforcement learning, which humans utilize in everyday life, as well as exploring when more sophisticated versions of neuromodulation like multiplexing and metamodulation are required is a fascinating direction of future research.

Our model of neuromodulation suggests several interesting connections to biological and clinical phenomena, but also has limitations. In addition to those noted above due to the simplified model we investigated, our study is rooted in artificial neural network (ANN) models. This approach has major advantages in studying fundamental principles of how neuromodulation of synaptic weights can modify parallel, distributed circuit computations, but also is strongly limited in its explanatory power of the full biological phenomena, as in psychiatric diseases, which have many more relevant parameters (e.g., neuron subtypes, circuit architecture, glia, environmental and historical factors, vasculature) not accounted for in our ANN models. Given these limitations of our model, our findings require experimental validation.

Finally, neuromodulation provides interesting directions for machine learning (ML). By separating synaptic memory regimes in a single network, we demonstrate how a network can have much greater flexibility and increased capacity, supporting a library of unique behaviors for overlapping external contingencies. Furthermore, each behavior can be rapidly accessed through targeted application of the relevant neuromodulatory factor. High-capacity, compact networks with high-speed access to different output modes presents a promising component for ML development and storage-limited applications like edge computing. Additionally, through the separation of memory regimes that effectively splits a single RNN into multiple processors, this mechanism may provide a means of realizing the super-Turing capability of specific RNN configurations as defined by Siegelmann (1995). The addition of a continuous modulating variable, the neuromodulation factor, may be sufficient to surpass Turing computation, analogous to Cabessa and Siegelman's (2014) finding that synaptic

plasticity enables super-Turing computation within RNNs. Future theoretical assessment of neuromodulated RNNs' capacity will establish if this simple mechanism is sufficient to exceed the Turing limit.

4 Methods

4.1 Positive-Negative Output Task. In the positive-negative output task two possible stimuli could be given to the agent (RNN in our simulations): either a positive pulse referred to as positive stimulus or +; or no pulse, also referred to as null stimulus or ∅ (see the extended data in the online supplement for details of input signals). The agent is trained to give a positive output (+1) for the positive stimulus and zero output (0) for the null stimulus. When a neuromodulation effect is applied (e.g., broadcast synaptic scaling by factor of 0.5), the agent is trained to output a different set of behaviors to the same stimuli: zero output for the positive stimulus and negative output for the null stimulus.

The positive-negative output task has many similarities to the well-known go-no go task which is a widely used experimental paradigm. The positive-negative output task differs in that there are two behavioral repertoires that can be elicited depending on the internal state of the agent that corresponds to the presence or absence of the neuromodulatory effect in our RNN models. The internal state typically transitions over longer timescales than stimulus-driven shifts in behavior; our task was designed to depict these longer timescale shifts in the internal state that drive diverse behaviors.

The three behavior and nine behavior variants of the positive-negative output task followed a similar paradigm with additional added behaviors. In the three-behavior version, a third behavior of positive stimulus → negative output, null stimulus → positive output was added. The nine-behavior version was similarly elaborated with nine possible behaviors. See the supplementary online file for details. In our simulations, each trial lasted 200 time steps of 5 ms each for a total represented duration of 1 second.

4.2 Neuromodulatory Neural Network Model. For our simulations, we used a continuous rate recurrent neural network (RNN) model with biologically plausible parameters similar to RNNs in prior works (Kim & Sejnowski, 2020; Trujillo et al., 2020). Consistent with biological neural networks, we implemented Dale's Law using the method in Song et al. (2016) such that each neuron was either excitatory or inhibitory. For all our simulations, we used an RNN with $N = 200$ neuron units, 80% excitatory and 20% inhibitory. In the RNN, each neuron can be connected to any other neuron with probability p_{con} (p_{con} was initialized at 0.8 in our simulations), and each neuron receives weighted input from connected neurons to produce a

firing rate governed by the neural dynamical equation,

$$\boldsymbol{\tau}\frac{d\boldsymbol{x}}{dt} = -\boldsymbol{x} + W\boldsymbol{r} + W_{in}u + N(0, 0.1), \tag{4.1}$$

where $\boldsymbol{\tau} \in \mathbb{R}^{1\times N}$ is the synaptic decay time constant for the N neurons in the network, $\boldsymbol{x} \in \mathbb{R}^{1\times N}$ is the synaptic current variable for the N neurons, $W \in \mathbb{R}^{N\times N}$ is the matrix of synaptic weights between all N neurons, $\boldsymbol{r} \in \mathbb{R}^{1\times N}$ is the output firing rates of the N neurons in the network, $W_{in} \in \mathbb{R}^{1\times N}$ are weights associated with external input u, and $N(0, 0.1)$ is added noise drawn from a normal distribution with mean 0 and variance 0.1. The output firing rate for the neurons is given by an element-wise nonlinear transfer function transformation of the synaptic current variable. In our network, we used the standard logistic sigmoid function as implemented by prior models (Kim & Sejnowski, 2020; Kim et al., 2019):

$$\boldsymbol{r} = \frac{1}{1 + e^{-x}}. \tag{4.2}$$

The synaptic connectivity matrix W was randomly initialized from a normal distribution with zero mean and standard deviation $g/\sqrt{N \cdot p_{con}}$, where g is the gain. We set $g = 1.5$ as previous studies have shown that networks operating in a high-gain regime ($g \geq 1.5$) support rich dynamics analogous to those of biological networks (Kim & Sejnowski, 2020; Kim et al., 2019; Laje & Buonomano, 2013). The synaptic decay time constants were randomly initialized to a value in the biologically plausible range of 20 to 100 ms. As in Kim et al. (2019), we used the first-order Euler approximation method to discretize equation 4.1 for the simulations; for neuron i:

$$x_{i,t} = \left(1 - \frac{\Delta t}{\tau}\right) x_{i,t-1} + \frac{\Delta t}{\tau}\left(\sum_j W_{ji} r_{ji,t-1} + W_{ui} u_{t-1}\right) + N\left(0, 0.1\right). \tag{4.3}$$

Output was generated by taking all recurrent network neurons' activities and passing them through a weighted output unit,

$$o_{network} = W_{out}\boldsymbol{r} + b_{out},$$

where $W_{out} \in \mathbb{R}^{N\times 1}$ are the neural output weights and b_{out} is the output unit's bias term. RNNs were trained by backpropagation through time using AdamOptimizer with a least root squared error objective function:

$$\mathcal{L} = \sqrt{\sum_{t=1}^{T}(z_t - o_t)^2},$$

where z_t is the desired output and o_t is the network output at time t in a trial of total length T ($T = 200$ time steps where each time step represents 5 ms, for a total trial length of 1 second in our simulations). For the positive-negative output task, desired output was a 200 time step output trace as described in section 4.1. For the *Drosophila* task, the desired output was a single target value as described in section 4.8; in this case, only one time step of output (the final network output at $t = 100$) contributed to the loss.

To apply an amplifying or dampening neuromodulatory effect, target neurons' weights were scaled by the neuromodulatory factor. For whole network neuromodulation, this effect was applied to all neurons in the RNN; for subpopulation neuromodulation, the effect was applied only to the selected subpopulation of neurons.

For the positive-negative output task, RNNs were trained until one of two possible stop criteria was met: (1) average trial root square error over the last 50 trials was under a threshold of 1, or (2) 10,000 training trials was reached. Performance on the task was then assessed by evaluating the percentage of test trials that matched the following performance criteria: for positive output trials, output was required to reach 1.0 ± 0.2 by time step 120 (full trial was 200 time steps); for zero output trails, output was required to be 0.0 ± 0.2 and for negative output trails -1.0 ± 0.2 at time step 120.

For the *Drosophila* task, RNNs were trained until one of two possible stop criteria was met: (1) average trial square error over the last 100 trials was under a threshold of 0.005 or (2) 10,000 training trials were reached. Output of RNNs was then assessed to determine network MAT (see mean acceptance threshold (MAT) description that follows) and behavior at inputs corresponding to standard sugar concentrations taken from Inagaki et al. (2012).

4.3 Comparison to Context-Dependent Cued Model. For comparison, we created a cue-driven RNN model and trained it on the positive-negative output task. The cue, which we refer to as the "context cue," was delivered to the RNN as an additional input stimulus and signaled which output behavior was desired. We created models with two types of cues: transient cues were constant value inputs present only during the stimulus input period ($t = 1$ to $t = 75$); persistent cues were constant value inputs across the whole trial. For comparisons between models, we ran models with cue pairs of $+1.0/-1.0$ (2.0 sep in Extended Data Figures 2a and 2b), $+0.5/-0.5$ (1.0 sep), and $+0.2/-0.2$ (0.4 sep).

We compared neuromodulatory and context-cued RNNs tolerance to noise in input signals and noise in neuromodulatory signal. For each noise condition we tested, we ran 100 simulations in four replicate RNNs and measured consistency of final network states using a measure of mean Euclidean distance. For details of noise simulations' comparison to context-dependent cued model, see the extended data in the online supplement.

4.4 Fixed-Point Analysis. Fixed-point analysis attempts to find stationary points in activity space that help characterize reproducible network dynamics between independently initialized and trained neural networks (Driscoll et al., 2024). We followed the fixed-point analysis described in Driscoll et al. (2024) and applied it to neuromodulated RNNs. To do this we trained RNNs with neuromodulation factor 2 on the positive-negative output task. After successful training, we used the fixed point finder package (https://github.com/mattgolub/fixed-point-finder; Golub & Sussillo, 2018) to search for stationary points corresponding to correct model output. We used the default arguments for fixed-point finder, following the example of the 3-bit flip-flop, including 5000 maximum iterations and 1000 initial states. We performed this analysis for five replicate RNNs independently initialized and trained, and for each replicate RNN, we ran the analysis both with neuromodulation present and absent. For q value cutoffs, we used $q << 1/T^2$ where T is the task timescale; for positive-negative output task $T = 200$, so $q << 2e\text{-}5$ for a fixed point similar to the q values used in Driscoll et al. and specifically tailored to our task time length. Minimum q values did not reach criteria of being sufficiently small to be considered fixed points (see Extended Data Figure 6). To confirm that the state vector had velocity while the model output remained sufficiently constant to meet our performance criteria for the task, we calculated the magnitude of the velocity vector over time from our hidden states and found this continued to fluctuate when output was constant and remained on the scale of $\sqrt{q}$. To evaluate if RNN dynamics were moving in a particular direction versus following more random motion, we measured the cosine similarity between velocity vectors at successive time points and found some directionally correlated movement during the beginning of the task followed by apparently random motion later on, possibly resulting from the noise term in the RNN's governing equation 4.1. While the hidden state vector does not, strictly speaking, settle into a stable fixed point, it does appear to fluctuate randomly within a relatively confined region of activity space. Given that the hidden state dimension is much greater than the output dimension, multiple hidden states map to similar enough output states to meet our performance criteria, which permits the network to wander through this space with stable output. Our RNNs' lack of fixed points while other RNNs consistently yield reproducible fixed points (Driscoll et al., 2024) likely is due to specific network design (i.e., governing equation including noise terms and activation functions), training procedures, hyperparameter selections, and the specific task the networks are trained on.

4.5 Neuromodulation of Multiple Subpopulations and Multiple Levels. For neuromodulation of nonoverlapping subpopulations, same-sized groups of neurons were chosen randomly without any overlap and neuromodulator applied to each for a given behavior. For overlapping subpopulations, groups of neurons were chosen randomly allowing overlap

(see Extended Data Figure 8). Our RNN models were initialized with random connections W as described above; thus, subpopulations of neurons selected randomly for neuromodulation targeting were not spatially localized. The subset of neuromodulated neurons in our models is analogous to a subset of neurons that receive neuromodulatory projections in the brain, which are often spatially dispersed subpopulations.

Neuromodulation at different levels ("multifactor networks") was done by applying different neuromodulation factors (f_{nm}). For the nine-behavior positive-negative output task, this was done using factors $\in [1:1:9]$; that is, for Behavior 1, no factor was applied ($f_{nm} = 1$); for Behavior 2, $f_{nm} = 2$; for Behavior 3, $f_{nm} = 3$, and so on.

To test networks across the range of subpopulation sizes with single or multiple neuromodulator factors on n-behavior positive-negative output tasks, stop criteria were adjusted to account for the increased behaviors: (1) average trial root square error over last n*25 trials was under threshold of 1, or (2) 15,000 training trials was reached. Performance on the tasks was assessed as before. For overlapping subpopulations, overlap was quantified in two ways. For each network, the number of neurons neuromodulated in two or more subpopulations was measured (see Extended Data Figure 8d). Overlap was also quantified by measuring the average number of neuromodulated subpopulations; a neuron in the network was a member of (see Extended Data Figure 8e).

4.6 Single Neuron Inputs, Functional Clustering, and Selectivity Index. Neural activity in an RNN is a complex function of all the neuron activities tracing all the way back in time. To understand how neuromodulation shifted synaptic inputs at the single-neuron level, we considered the first time point in a trial. For any trial, at $t = 0$ all activities are randomly initialized from a normal distribution. As a result, at $t = 1$, a neuron reacts only to the weighted inputs of its incoming connections, uncontaminated by propagating recurrent activity dynamics from past time points. Analysis of neuron activity at this time point is shown in Extended Data Figures 9a to 9c.

In order to examine whether trained models contained functionally specialized neurons, we grouped neuron activities by combination of subtask (which maps one-to-one with neuromodulatory state) and stimulus given (stimulus-subtask combinations). We averaged activity of each neuron over time and trials within each group. This resulted in a matrix of time-trial averaged neuron activities with a number of rows equal to the stimulus-subtask combinations, and number of columns equal to the number of neurons. Using k-means, we clustered neurons with similar activity levels across stimulus-subtask combinations. We computed a silhouette score to find the optimal number of clusters, which for the RNN in Extended Data Figure 9, was 6. The silhouette score computed for five and six clusters differed only by 0.7%, and the additional cluster was very small and similar to

an existing cluster, so we conducted further analysis with five clusters for simplicity.

To measure the selectivity of individual neurons for particular stimulus-subtask combinations, we calculated a selectivity index (si) for each neuron j:

$$\mathrm{si}_j = \frac{\bar{r}_j^{\max} - \bar{r}_j^{second\_\max}}{\bar{r}_j^{\max}},$$

where $\bar{r}_j$ is the average firing rate of neuron j over the trial duration, $\bar{r}_j^{max}$ indicates the maximum $\bar{r}_j$ across all the stimulus-subtask combinations and $\bar{r}_j^{second\_max}$ indicates the second highest $\bar{r}_j$ over all the stimulus-subtask combinations. The selectivity index thus captures a normalized approximation of how uniquely active a neuron was for a given stimulus-subtask combination.

4.7 Network Population Dynamics. To represent whole network population activity dynamics, we sought a low-dimensional representation of whole population activity. We used principal component analysis (PCA) since the leading components capture the largest projections of activity variability, which we hypothesized would effectively separate our neuromodulatory conditions if large differences occurred (Cunningham & Yu, 2014). We found this was the case. We found qualitatively similar results using multidimensional scaling, which finds projections designed to best preserve distances in high-dimensional activity space. For our figures, we display the first three PCs, as these captured a large amount of the activity variance (80%–92% explained across the analyses) and effectively represent the activity dynamic differences in the analyses.

To map the neuromodulation-dependent activity subspace, we generated 100 independent stimuli series consisting of random numbers drawn from a uniform distribution between 0 and 1 at each time point ($t = 0$ to $t = 200$) and fed this into the RNN with and without neuromodulation (shotgun stimulus mapping from Extended Data Figure 12a), which defined nonoverlapping subspaces of activity space.

To analyze neuromodulation transition curves, we compared activity under intermediate neuromodulation levels with linear interpolation. Linear interpolation was done by evenly dividing the distance between no and full neuromodulation activity states into nine sections analogous to the nine neuromodulation levels assessed. These nine points in activity space were then used to generate output that is plotted in Extended Data Figure 13. The geometry of the neuromodulation-based transition was assessed by calculating the Euclidean distance of intermediate neuromodulation-level states at a given trial time point to the nearest point on the line connecting no and full neuromodulation states at that time point. These distances

are plotted in Figure 3i. The angle of departure (AoD) was defined as the angle formed by the line between no and full neuromodulation states and the line between no neuromodulation and the first neuromodulation level states, which can be calculated as

$$\vec{v_1} = \vec{p_F} - \vec{p_N},$$
$$\vec{u_1} = \vec{p_{L1}} - \vec{p_N},$$
$$AoD = cos^{-1} \frac{\vec{u_1} \cdot \vec{v_1}}{|\vec{u_1}||\vec{v_1}|},$$

where $\vec{p_N}$ is the network state with no neuromodulation, $\vec{p_F}$ is the network state with full neuromodulation, and $\vec{p_{L1}}$ is the network state with the first level of neuromodulation.

4.8 EC50. The EC50 of a network was defined as the level of neuromodulation that led to half the output of full neuromodulation, analogous to EC50 measured in dose-response curves for pharmaceuticals: the dose ("effective concentration") that drives a half-maximal (50%) response. For the results reported, we used EC50 calculated for the positive stimulus. For this stimulus in the positive-negative output task, a nonneuromodulated network outputs +1 and a fully neuromodulated network outputs 0 (measurements for output level were taken at 0.5 s through the trial); the EC50 for the network in this case is the amount of neuromodulator required to output 0.5. The EC50 was calculated by fitting a sigmoid curve to the progression of output (from +1 to 0 in this case) with increasing the neuromodulation level (see Figure 3d),

$$output = 1 - \frac{1}{1 + e^{a \cdot f_{nm} + b}},$$

where f_{nm} is the neuromodulation level. The EC50 neuromodulation level was calculated by finding the intersection of the sigmoid and the half-maximal output; for half-maximal output of 0.5, EC50 $= -b/a$. Sigmoid curves were fit using a least squares fit.

To assess the variability of individual network sensitivities to intermediate levels of neuromodulation, 29 RNNs were independently trained with a neuromodulation factor of 9. Since our focus in this analysis was on the variability of sensitivity to neuromodulator exhibited by networks, approximately 30 independent samples of RNNs was sufficient to show the high variability: a 3.1 × spread, ranging from 2.1 to 6.5 with a full neuromodulation factor of 9. This demonstrated that occurrence of low, and high-sensitivity networks occurred with relatively high likelihood—roughly at least 1/30 or 3%.

4.9 *Drosophila* Sugar Sensitivity Task. We implemented a computational version of Inagaki et al. (2012) to train our network models. During training, models were presented with a constant sugar concentration (external input proportional to sugar concentration) for 100 time steps (equivalent to 500 ms) and trained to output a probability of proboscis extension reflex (PER), a behavior exhibited by *Drosophila* where the proboscis is extended in response to a desirable food stimulus. This network output represents the upstream signal of executing the PER behavior; thus, when a network output was 50%, 5 times out of 10 that network would do the PER. This directly would translate into the fraction of flies on average that would exhibit PER in that condition as measured in Inagaki et al. (2012). Curves in Figure 4b show average behavior of multiple networks, where each network showed unique behavior transitions as reflected in Extended Data Figure 14d. For fed and 2-day starved training, we used a piece-wise linear approximation estimated from Inagaki et al. (2012) to determine desired output values. To compare box plots of MAT, one-way ANOVA followed by *t*-test with Bonferroni correction was used as in Inagaki et al. (2012).

During training, networks received an input stimulus corresponding to a sugar concentration drawn from a continuous uniform distribution between a maximum and minimum sugar concentration (6.25 mM and 800 mM as in Inagaki et al., 2012). Networks were trained to output the sugar sensitivity curves corresponding to fed (no neuromodulation) and 2-day starved (full neuromodulation, factor 5). After training, networks were tested with no neuromodulation, full neuromodulation (factor 5), and an intermediate neuromodulation level not seen during training (factor 3).

4.10 Mean Acceptance Threshold (MAT). MAT is the sugar concentration at which $PER = 0.5$. Analogous to the analysis done for flies in Inagaki et al. (2012), for each RNN a sigmoid was fit,

$$PER = \frac{1}{1 + e^{-a \cdot \log_2 \frac{x_{sugar}}{MAT}}},$$

where a is the slope of the sigmoid. When $PER = 0.5$, then $x_{sugar} = MAT$. Sigmoid curves were fit using a least squares fit.

For intermediate neuromodulatory level ($f_{nm} = 3$) MAT variability analysis, a normalized change in MAT (%ΔMAT) was calculated:

$$\%\Delta MAT = \frac{MAT_{f_{nm}=3} - MAT_{f_{nm}=1}}{MAT_{f_{nm}=5} - MAT_{f_{nm}=1}},$$

where $MAT_{f_{nm}=x}$ is the RNN's MAT with neuromodulation factor x. %ΔMAT gives a network normalized metric for how much the intermediate neuromodulation (f_{nm}=3) moved the fly from no neuromodulation (f_{nm}=1) to full neuromodulation (f_{nm}=5) sensitivity.

4.11 Electrical Modulation. We administered electrical modulation as an external current applied for the duration of the trial. "On-target" modulation was applied to the neuromodulated neuron population, and "random" modulation was applied to a randomly selected group of neurons of equal size; these could include both neuromodulated or nonneuromodulated neurons. All neurons (both excitatory and inhibitory) within the selected subpopulation were given identical external current modulation. For fixed electrical modulation simulations (see Figures 5b and Extended Data Figures 15a and 15b), a current of magnitude 1 was applied (+1 for excitatory modulation; −1 for inhibitory).

The initial experiment testing the e-mod effect (see Figure 5b and Extended Data Figures 15a and 15b) was done in 10 independently trained RNNs with 50% subpopulation of neurons neuromodulated by a factor of 0.5. After identifying the effect of e-mod on RNNs, we sought to compare e-mod sensitivity to neuromodulation sensitivity. To do this, we trained 30 RNNs with 50% subpopulation neuromodulated and a factor of 9 (similar to the sensitivity analysis in Figure 3). We then administered graded electrical modulation (e-mod $\in -\{0.0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0\}$) and assessed network sensitivity (e-mod50) the same way as for neuromodulation (EC50). For these networks, we also measured neuromodulation sensitivity for comparison. These results are shown in Figures 5c to 5f and Extended Data Figure 15c. For details see the online supplementary material.

Test performance (Figures 5b and Extended Data Figures 15a and 15b) was measured in the same way as described previously for the positive-negative output task by assessing the percentage of test trials that matched the following performance criteria: for positive output trials, output was required to reach 1.0 ± 0.2 by time step 120 (full trial was 200 time steps); for zero output trials, output was required to be 0.0 ± 0.2 and for negative output trails -1.0 ± 0.2 at time step 120. To assess decreases in test performance shown in Figure 5b, further analysis was done to characterize the breakdown of specific behaviors as shown in Extended Data Figures 15a and 15b.

Though similar to exogenous cues (Mante et al., 2013) in the mode of delivery (compare $W_{ui}u$ and u_{Estim} terms in the last equation of appendix B in the Extended Data), electrical modulation in these experiments was delivered to an RNN after it had been trained rather than during training when the network could learn to use it as a signal. In this way, our implementation of electrical modulation is more akin to experimental use of optogenetics in probing neuromodulation-dependent network behavior or clinical administration of DBS and TMS in therapeutic modulation.

5 Conclusion

Neuromodulation, through the release of molecules like serotonin, norepinephrine, and dopamine, provides a control mechanism that allows

brains to shift into distinct behavioral modes. We used an artificial neural network model to show how neuromodulatory molecules acting as a broadcast signal on synaptic connections enables flexible and smooth behavioral shifting. We find that individual networks exhibit idiosyncratic sensitivities to neuromodulation under identical training conditions, highlighting a principle underlying behavioral variability. Network sensitivity is tied to the geometry of network activity dynamics, which provides an explanation for why different types of neuromodulation (e.g., molecular versus direct current modulation) have different behavioral effects. Our work suggests experiments to test biological hypotheses and paths forward in the development of flexible artificial intelligence systems.

Acknowledgments

We thank Yusi Chen and Robert Kim for helpful discussions, Ramona Marchand for project support, and Jorge Aldana for support with computing resources. Research was supported by the Kavli Institute for Brain and Mind at UC San Diego, ONR grant N00014-16-1-2829, DARPA grant HR0011-18-2-002, NIMH grant R37-MH102441, and National Center for Complementary and Integrative Health Pioneer Award DP1-AT009925.

Author Contributions

B.T., S.C.P., K.M.T., H.T.S., and T.J.S. formulated the ideas; B.T., S.C.P. and T.J.S. designed the network architecture and simulation of the behavioral tasks; B.T. and S.C.P. implemented the design, ran the simulations, and carried out the data analysis; B.T., S.C.P, K.M.T., H.T.S., and T.J.S. wrote the manuscript.

Declaration of Interests

The authors declare no competing interests.

References

Allen, W. E., Chen, M. Z., Pichamoorthy, N., Tien, R. H., Pachitariu, M., Luo, L., & Deisseroth, K. (2019). Thirst regulates motivated behavior through modulation of brainwide neural population dynamics. *Science*, *364*, eaav3932. 10.1126/science.aax6175

Alonso, L. M., & Marder, E. (2020). Temperature compensation in a small rhythmic circuit. *eLife*, *9*, e55470. 10.7554/eLife.55470

Ashok, A. H., Marques, T. R., Jauhar, S., Nour, M. M., Goodwin, G. M., Young, A. H., & Howes, O. D. (2017). The dopamine hypothesis of bipolar affective disorder: The state of the art and implications for treatment. *Molecular Psychiatry*, *22*, 666–679. 10.1038/mp.2017.16

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450. 10.1146/annurev.neuro.28.061604.135709

Avery, M. C., & Krichmar, J. L. (2017). Neuromodulatory systems and their interactions: A review of models, theories, and experiments. *Frontiers in Neural Circuits*, *11*, 108. 10.3389/fncir.2017.00108

Bacqué-Cazenave, J., Bharatiya, R., Barrière, G., Delbecque, J., Bouguiyoud, N., Giovanni, G. D., Cattaert, D., & Deurwaerdère, P. D. (2020). Serotonin in animal cognition and behavior. *International Journal of Molecular Sciences*, *21*, 1649.

Bargmann, C. I. (2012). Beyond the connectome: How neuromodulators shape neural circuits. *Bioessays*, *34*, 458–465. 10.1002/bies.201100185

Beaulieu, S., Frati, L., Miconi, T., Lehman, J., Stanley, K. O., Clune, J., & Cheney, N. (2020). Learning to continually learn. In G. D. Giacomo, A. Catalá B. Dilkina, M. Milano, S. Barro, A. Bugarín, & J. Lang (Eds.), *24th European Conference on Artificial Intelligence* (pp. 992–1001).

Blokhin, I. O., Khorkovad, O., Saveanub, R. V., & Wahlestedt, C. (2020). Molecular mechanisms of psychiatric diseases. *Neurobiology of Disease*, *146*, 105136. 10.1016/j.nbd.2020.105136

Burgos-Robles, A., Kimchi, E. Y., Izadmehr, E. M., Porzenheim, M. J., Ramos-Guasp, W. A., Nieh, E. H., . . . Tye, K. M. (2017). Amygdala inputs to prefrontal cortex guide behavior amid conflicting cues of reward and punishment. *Nature Neuroscience*, *20*, 824–835. 10.1038/nn.4553

Cabessa, J., & Siegelmann, H. T. (2014). The super-Turing computational power of plastic recurrent neural networks. *International Journal of Neural Systems*, *24*(8), 1450029. 10.1142/S0129065714500294

Calhoun, A. J., Pillow, J. W., & Murthy, M. (2019). Unsupervised identification of the internal states that shape natural behavior. *Nature Neuroscience*, *22*, 2040–2049. 10.1038/s41593-019-0533-x

Cermak, N., Yu, S. K., Clark, R., Huang, Y.-C., Baskoylu, S. N., & Flavell, S. W. (2020). Whole-organism behavioral profiling reveals a role for dopamine in state-dependent motor program coupling in *C. elegans*. *eLife*, *9*, e57093. 10.7554/eLife.57093

Cunningham, J. P., & Yu, B. M. (2014). Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*, *17*, 1500–1508. 10.1038/nn.3776

Dayan, P. (2012). Twenty-five lessons from computational neuromodulation. *Neuron*, *76*, 240–256. 10.1016/j.neuron.2012.09.027

Deisseroth, K. (2015). Optogenetics: 10 years of microbial opsins in neuroscience. *Nature Neuroscience*, *18*, 1213–1225. 10.1038/nn.4091

Dethier, J., Drion, G., Franci, A., & Sepulchre, R. (2015). A positive feedback at the cellular level promotes robustness and modulation at the circuit level. *Journal of Neurophysiology*, *114*, 2472–2484. 10.1152/jn.00471.2015

Disney, A. A. (2021). Neuromodulatory control of early visual processing in macaque. *Annual Review of Vision Science*, *7*, 181–199. 10.1146/annurev-vision-100119-125739

Driscoll, L. N., Shenoy, K., & Sussillo, D. (2024). Flexible multitask computation in recurrent networks utilizes shared dynamical motifs. *Nature Neuroscience*, *27*, 1349–1363. 10.1038/s41593-024-01668-6

Fellous, J.-M., & Linster, C. (1998). Computational models of neuromodulation. *Neural Computation*, *10*, 771–805. 10.1162/089976698300017476

Furukawa, T. A., Cipriani, A., Cowen, P. J., Leucht, S., Egger, M., & Salanti, G. (2019). Optimal dose of selective serotonin reuptake inhibitors, venlafaxine, and mirtazapine in major depression: A systematic review and dose-response meta-analysis. *Lancet Psychiatry*, *6*, 601–609. 10.1016/S2215-0366(19)30217-2

Goldman, M. S., Golowasch, J., Marder, E., & Abbott, L. F. (2001). Global structure, robustness, and modulation of neuronal models. *Journal of Neuroscience*, *21*, 5229–5238. 10.1523/JNEUROSCI.21-14-05229.2001

Golub, M. D., & Sussillo, D. (2018). Fixedpointfinder: A tensorflow toolbox for identifying and characterizing fixed points in recurrent neural networks. *Journal of Open Source Software*, *3*(31), 1003. 10.21105/joss.01003

Gorfine, T., Assaf, Y., Goshen-Gottstein, Y., Yeshurun, Y., & Zisapel, N. (2006). Sleep-anticipating effects of melatonin in the human brain. *NeuroImage*, *31*, 410–418. 10.1016/j.neuroimage.2005.11.024

Goudar, V., & Buonomano, D. V. (2018). Encoding sensory and motor patterns as time-invariant trajectories in recurrent neural networks. *eLife*, *7*, e31134. 10.7554/eLife.31134

Gouveia, F. V., Gidyk, D. C., Giacobbe, P., Ng, E., Meng, Y., Davidson, B., . . . Hamani, C. (2019). Neuromodulation strategies in post-traumatic stress disorder: From preclinical models to clinical applications. *Brain Sciences*, *9*, 45. 10.3390/brainsci9020045

Gutierrez, G. J., & Marder, E. (2013). Rectifying electrical synapses can affect the influence of synaptic modulation on output pattern robustness. *Journal of Neuroscience*, *33*, 13238–13248. 10.1523/JNEUROSCI.0937-13.2013

Haddad, S. A., & Marder, E. (2018). Circuit robustness to temperature perturbation is altered by neuromodulators. *Neuron*, *100*(3), 609–623. 10.1016/j.neuron.2018.08.035

Harris-Warrick, R. M., & Johnson, B. R. (2010). Checks and balances in neuromodulation. *Frontiers in Behavioral Neuroscience*, *4*, 47.

Hasselmo, M. E. (2006). The role of acetylcholine in learning and memory. *Current Opinion in Neurobiology*, *16*, 710–715. 10.1016/j.conb.2006.09.002

Hasselmo, M. E., Schnell, E., & Barkai, E. (1995). Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience*, *15*, 5249–5262. 10.1523/JNEUROSCI.15-07-05249.1995

Hyman, S. E. (2013). Psychiatric drug development: Diagnosing a crisis. *Cerebrum*, *5*.

Inagaki, H. K., de-Leon, S. B.-T., Wong, A. M., Jagadish, S., Ishimoto, H., Barnea, G., . . . Anderson, D. J. (2012). Visualizing neuromodulation in vivo: Tango-mapping of dopamine signaling reveals appetite control of sugar sensing. *Cell*, *148*, 583–595.

Jang, H., Kim, K., Neal, S. J., Macosko, E., Kim, D., Butcher, R. A., . . . Sengupta, P. (2012). Neuromodulatory state and sex specify alternative behaviors through antagonistic synaptic pathways in *C. elegans*. *Neuron*, *75*, 585–592. 10.1016/j.neuron.2012.06.034

Jing, J., Gillette, R., & Weiss, K. R. (2009). Evolving concepts of arousal: Insights from simple model systems. *Reviews in the Neurosciences*, *20*, 405–427. 10.1515/REVNEURO.2009.20.5-6.405

Johnson, B. R., Brown, J. M., Kvarta, M. D., Lu, J. Y. J., Schneider, L. R., Nadim, F., & Harris-Warrick, R. M. (2011). Differential modulation of synaptic strength and timing regulate synaptic efficacy in a motor network. *Journal of Neurophysiology*, *105*, 293–304. 10.1152/jn.00809.2010

Johnson, B. R., Peck, J. H., & Harris-Warrick, R. M. (1995). Distributed amine modulation of graded chemical transmission in the pyloric network of the lobster stomatogastric ganglion. *Journal of Neurophysiology*, *74*, 437–452. 10.1152/jn.1995.74.1.437

Katz, P. S., & Edwards, D. H. (1999). Metamodulation: The control and modulation of neuromodulation. In P. S. Katz (Ed.), *Beyond neurotransmission: Neuromodulation and its importance for information processing* (pp. 349–381). Oxford University Press.

Katz, P. S., Getting, P. A., & Frost, W. N. (1994). Dynamic neuromodulation of synaptic strength intrinsic to a central pattern generator circuit. *Nature*, *367*, 729–731. 10.1038/367729a0

Kim, R., Li, Y., & Sejnowski, T. J. (2019). Simple framework for constructing functional spiking recurrent neural networks. *PNAS*, *116*(45), 22811–22820. 10.1073/pnas.1905926116

Kim, R., & Sejnowski, T. J. (2020). Strong inhibitory signaling underlies stable temporal dynamics and working memory in spiking neural networks. *Nature Neuroscience*, *24*.

Krauss, J. K., Lipsman, N., Aziz, T., Boutet, A., Brown, P., Chang, J. W., . . . Lozano, A. M. (2021). Technology of deep brain stimulation: Current status and future directions. *Nature Reviews Neurology*, *17*, 75–87. 10.1038/s41582-020-00426-z

Laje, R., & Buonomano, D. V. (2013). Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nature Neuroscience*, *16*(7), 925–933. 10.1038/nn.3405

Lauschke, V. M., Zhou, Y., & Ingelman-Sundberg, M. (2019). Novel genetic and epigenetic factors of importance for inter-individual differences in drug disposition, response and toxicity. *Pharmacology Therapeutics*, *197*, 122–152. 10.1016/j.pharmthera.2019.01.002

Lee, S.-H., & Dan, Y. (2012). Neuromodulation of brain states. *Neuron*, *76*, 209–222. 10.1016/j.neuron.2012.09.012

Li, X., Bucher, D., & Nadim, F. (2018). Distinct co-modulation rules of synapses and voltage-gated currents coordinate interactions of multiple neuromodulators. *Journal of Neuroscience*, *38*, 8549–8562. 10.1523/JNEUROSCI.1117-18.2018

Low, R. J., Lewallen, S., Aronov, D., Nevers, R., & Tank, D. W. (2018). *Probing variability in a cognitive map using manifold inference from neural dynamics*. bioRxiv. 10.1101/418939

Maheswaranathan, N., Williams, A. H., Golub, M. D., Ganguli, S., & Sussillo, D. (2019). Universality and individuality in neural dynamics across large populations of recurrent networks. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in neural information processing systems*, *32*. Curran.

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, *503*, 78–84. 10.1038/nature12742

Marder, E. (2012). Neuromodulation of neuronal circuits: Back to the future. *Neuron*, *76*, 1–11. 10.1016/j.neuron.2012.09.010

Marder, E., & Eisen, J. S. (1984). Electrically coupled pacemaker neurons respond differently to same physiological inputs and neurotransmitters. *Journal of Neurophysiology*, *51*, 1362–1374. 10.1152/jn.1984.51.6.1362

Marder, E., & Thirumalai, V. (2002). Cellular, synaptic and network effects of neuromodulation. *Neural Networks*, *15*, 479–493. 10.1016/S0893-6080(02)00043-6

McCormick, D. A., Nestvogel, D. B., & He, B. J. (2020). Neuromodulation of brain state and behavior. *Annual Review of Neuroscience*, *43*, 391–415. 10.1146/annurev-neuro-100219-105424

Mehrer, J., Spoerer, C. J., Kriegeskorte, N., & Kietzmann, T. C. (2020). Individual differences among deep neural network models. *Nature Communications*, *11*, 5725. 10.1038/s41467-020-19632-w

Miconi, T., Rawal, A., Clune, J., & Stanley, K. O. (2019). Backpropamine: Training self-modifying neural networks with differentiable neuromodulated plasticity. In *Proceedings of the 7th International Conference on Learning Representations.*

Nadim, F., & Bucher, D. (2012). Neuromodulation of neurons and synapses. *Current Opinion in Neurobiology*, *29*, 48–56. 10.1016/j.conb.2014.05.003

Newman, E. L., Gupta, K., Climer, J. R., Monaghan, C. K., & Hasselmo, M. E. (2012). Cholinergic modulation of cognitive processing: Insights drawn from computational models. *Frontiers in Behavioral Neuroscience*, *6*, 24. 10.3389/fnbeh.2012.00024

Nieh, E. H., Schottdorf, M., Freeman, N. W., Low, R. J., Lewallen, S., Koay, S. A., . . . Tank, D. W. (2021). Geometry of abstract learned knowledge in the hippocampus. *Nature*, *595*(7865), 80–84. 10.1038/s41586-021-03652-7

O'Callaghan, C., Walpola, I. C., & Shine, J. M. (2020). Neuromodulation of the mind-wandering brain state: The interaction between neuromodulatory tone, sharp wave-ripples and spontaneous thought. *Philosophical Transactions of the Royal Society B*, *376*, 20190699.

Ozomaro, U., Wahlestedt, C., & Nemeroff, C. B. (2013). Personalized medicine in psychiatry: Problems and promises. *BMC Medicine*, *11*, 132. 10.1186/1741-7015-11-132

Patriarchi, T., Cho, J. R., Merten, K., Howe, M. W., Marley, A., Xiong, W.-H., . . . Tian, L. (2018). Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science*, *360*, eaat4422. 10.1126/science.aat4422

Prinz, A. A., Bucher, D., & Marder, E. (2004). Similar network activity from disparate circuit parameters. *Nature Neuroscience*, *7*, 1345–1352. 10.1038/nn1352

Provenza, N. R., Matteson, E. R., Allawala, A. B., Barrios-Anderson, A., Sheth, S. A., Viswanathan, A., . . . Borton, D. A. (2019). The case for adaptive neu-romodulation to treat severe intractable mental disorders. *Frontiers in Neuroscience*, *13*, 152. 10.3389/fnins.2019.00152

Radnikow, G., & Feldmeyer, D. (2018). Layer- and cell type-specific modulation of excitatory neuronal activity in the neocortex. *Frontiers in Neuroanatomy*, *12*. 10.3389/fnana.2018.00001

Reynolds, G. P., McGowan, O. O., & Dalton, C. F. (2013). Pharmacogenomics in psychiatry: The relevance of receptor and transporter polymorphisms. *British Journal of Clinical Pharmacology*, *77*, 654–672. 10.1111/bcp.12312

Ribeiro, J. A., & Sebastiao, A. M. (2010). Modulation and metamodulation of synapses by adenosine. *Acta Physiologica*, *199*, 161–169. 10.1111/j.1748-1716.2010.02115.x

Risio, L. D., Borgi, M., Pettorruso, M., Miuli, A., Ottomana, A. M., Sociali, A., . . . Zoratto, F. (2020). Recovering from depression with repetitive transcranial magnetic stimulation (rTMS): A systematic review and meta-analysis of preclinical studies. *Translational Psychiatry*, *10*, 393. 10.1038/s41398-020-01055-2

Russo, A. A., Bittner, S. R., Perkins, S. M., Seely, J. S., London, B. M., Lara, A. H., Miri, A., . . . Churchland, M. M. (2018). Motor cortex embeds muscle-like commands in an untangled population response. *Neuron*, *97*, 953–966. 10.1016/j.neuron.2018.01.004

Siegelmann, H. T. (1995). Computational beyond the Turing limit. *Science*, *268*, 545–548. 10.1126/science.268.5210.545

Song, H. F., Yang, G. R., & Wang, X.-J. (2016). Training excitatory-inhibitory recurrent neural networks for cognitive tasks: A simple and flexible framework. *PLOS Computational Biology*, *12*, 1–30.

Staudt, M. D., Herring, E. Z., Gao, K., Miller, J. P., & Sweet, J. A. (2019). Evolution in the treatment of psychiatric disorders: From psychosurgery to psychopharmacology to neuromodulation. *Frontiers in Neuroscience*, *13*, 108. 10.3389/fnins.2019.00108

Stroud, J. P., Porter, M. A., Hennequin, G., & Vogels, T. P. (2018). Motor primitives in space and time via targeted gain modulation in cortical networks. *Nature Neuroscience*, *21*, 1774–1783. 10.1038/s41593-018-0276-0

Tarder-Stoll, H., Jayakumar, M., Dimsdale-Zucker, H. R., Günseli, E., & Aly, M. (2020). Dynamic internal states shape memory retrieval. *Neuropsychologia*, *138*, 197328. 10.1016/j.neuropsychologia.2019.107328

Thair, H., Holloway, A. L., Newport, R., & Smith, A. D. (2017). Transcranial direct current stimulation (TDCS): A beginner's guide for design and implementation. *Frontiers in Neuroscience*, *11*. 10.3389/fnins.2017.00641

Trujillo, C. A., Adams, J. W., Negraes, P. D., Carromeu, C., Tejwani, L., Acab, A., . . . Muotri, A. R. (2020). Pharmacological reversal of synaptic and network pathology in human MECP2-KO neurons and cortical organoids. *EMBO Molecular Medicine*, *13*, e12523. 10.15252/emmm.202012523

Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J., & Aston-Jones, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science*, *283*, 549–554. 10.1126/science.283.5401.549

van Westrhenen, R., Aitchison, K. J., Ingelman-Sundberg, M., & Jukić, M. M. (2020). Pharmacogenomics of antidepressant and antipsychotic treatment: How far have we got and where are we going? *Frontiers in Psychiatry*, *11*, 94. 10.3389/fpsyt.2020.00094

Vecoven, N., Ernst, D., Wehenkel, A., & Drion, G. (2020). Introducing neuromodulation in deep neural networks to learn adaptive behaviors. *PLOS One*, *15*, e0227922. 10.1371/journal.pone.0227922

Voineskos, D., Daskalakis, Z. J., & Blumberger, D. M. (2020). Management of treatment-resistant depression: Challenges and strategies. *Neuropsychiatric Disease and Treatment*, *16*, 221–234. 10.2147/NDT.S198774

Weele, C. M. V., Siciliano, C. A., Matthews, G. A., Namburi, P., Izadmehr, E. M., Espinel, . . . Tye, K. M. (2018). Dopamine enhances signal-to-noise ratio in cortical-brainstem encoding of aversive stimuli. *Nature*, *563*, 397–401. 10.1038/s41586-018-0682-1

Wester, J. C., & McBain, C. J. (2014). Behavioral state-dependent modulation of distinct interneuron subtypes and consequences for circuit function. *Current Opinion in Neurobiology*, *29*, 118–125. 10.1016/j.conb.2014.07.007

Xie, P., Liu, L., Kuang, X., Wang, Y., Qu, L., Gong, H., Zeng, H. (2021). Morphological diversity of single neurons in molecularly defined cell types. *Nature*, *598*, 174–181. 10.1038/s41586-021-03941-1

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*, 681–692. 10.1016/j.neuron.2005.04.026

Zhang, S. X., Lutas, A., Yang, S., Diaz, A., Fluhr, H., Nagel, G., . . . Andermann, M. L. (2021). Hypothalamic dopamine neurons motivate mating through persistent camp signalling. *Nature*, *597*, 245–249. 10.1038/s41586-021-03845-0

Received September 3, 2025; accepted September 30, 2025.