A Appendix

A.1 Glossary

Backpropagation through time: An algorithm for training recurrent neural networks by unrolling the network over time steps and applying the standard backpropagation algorithm, allowing the network to learn from sequences of inputs and capture temporal dependencies.

Binary Cross-Entropy (BCE): A loss function commonly used in binary classification problems that measures the difference between predicted probabilities and actual binary labels, encouraging the model to output confident and accurate predictions.

Connectome: A comprehensive map of neural connections in the brain, detailing the structural connectivity between neurons or brain regions, providing insights into brain organization and function.

Covariance-based learning rule: A learning rule that updates weights based on the covariance between neural activity and reward, adjusting connection strengths to capture statistical relationships in the data and potentially improving the network's performance.

Deviance explained: A statistical measure of goodness of fit for generalized linear models, quantifying the proportion of variability in the data that is accounted for by the model compared to a null model.

Gaussian process: A statistical model used for regression and probabilistic classification, defining a distribution over functions and allowing for flexible, nonparametric modeling with uncertainty quantification.

Generative adversarial network (GAN) approach: A machine learning framework where two neural networks compete to generate realistic data, with one network (the generator) creating synthetic samples and another (the discriminator) trying to distinguish between real and fake data.

Homeostatic adjustments: Changes in neural systems that help maintain stability, including mechanisms like synaptic scaling and intrinsic plasticity that regulate neuronal activity and prevent excessive excitation or inhibition.

Logistic regression: A statistical method for predicting a binary outcome by modeling the probability of an event occurring as a function of input variables, using the logistic function to transform linear combinations of features into probabilities.

Synaptic plasticity: The ability of synapses (connections between neurons) to change their strength over time, forming the basis for learning and memory in the brain.

Truncated Taylor series: A mathematical method for approximating functions using polynomial terms, where the series is cut off after a finite number of terms to balance accuracy and computational tractability.

Meta-learning: A subfield of machine learning focused on improving the learning process itself, developing algorithms that can learn how to learn and adapt quickly to new tasks or environments.

Metaplasticity: Higher-order plasticity where prior synaptic activity influences subsequent plasticity, regulating the threshold and magnitude of future synaptic changes to maintain network stability and optimize learning.

Multilayer perceptrons (MLPs): A type of artificial neural network with multiple layers of nodes, including input, hidden, and output layers, capable of learning complex non-linear relationships in data through backpropagation.

Mushroom body: A region in the insect brain involved in learning and memory, particularly important for olfactory learning, sensory integration, and decision-making processes.

Neuromorphic: Referring to artificial systems that mimic biological neural systems, often implemented in hardware to achieve brain-like computation with

A.3 Additional detail: Inferring a plasticity rule from neural activity

To create a ground-truth system with a known plasticity rule, we use a single-layer neural network with 100 input neurons and $N = 1000$ output neurons to generate synthetic data. The output layer employs a sigmoid activation function, chosen for its differentiability and biologically plausible output range of [0, 1]. We generate 50 training trajectories for the network. The input data is sampled from a Gaussian distribution with zero mean and a variance of 0.1, independent of time. A subset of neurons, determined by the sparsity factor, is selected for the readout. The simulation runs over 50 time steps, calculating the neural activity of all output neurons at each step. To ensure numerical stability and prevent exploding gradients, gradient clipping is applied with a threshold of 0.2.

The coefficients $\theta_{\alpha\beta\gamma}$ of the Taylor series expansion representing the plasticity rule are learned, initialized independently and identically distributed (i.i.d.) from a normal distribution with a mean of 0 and a variance of 10^{-4}. Both the ground-truth and model network weights are initialized using Kaiming initialization from a zero-mean Gaussian distribution. Although these weights are drawn from the same distribution, they are resampled, resulting in different initial values. No regularization is applied during training. The Adam optimizer is used to train the weights of the plasticity model, with default parameters.

A.4 Additional detail: Inferring plasticity rules from behavior

In the behavioral simulation experiments, the ground truth plasticity rule is denoted as $x_j r$. We use a network with a 2-10-1 architecture and a sigmoid nonlinearity. The plasticity MLP has a size of 4-10-1. The default L1 regularization is set to 1e-2, the moving average window is 10, and the input firing mean is 0.75.

Given the discrete nature of the observed behavior and the continuous output of the model, we employ the percent deviance explained as a performance metric. The percent deviance explained measures the model's ability to account for the variability in observed binary choices compared to a null model that assumes no plasticity (i.e., the weights remain at their initial random initialization). It represents the reduction in deviance relative to this null model, expressed as a percentage. Higher values indicate greater log-likelihoods, signifying a superior fit to the observed data.

$$\text{Percent Deviance Explained} = 100 \times \left(1 - \frac{\text{Deviance}_{\text{model}}}{\text{Deviance}_{\text{null}}}\right) \tag{1}$$

Since there are 2 odors, they are encoded in a stimulus vector of dimension 2. Odor 1 corresponds to the first dimension "firing"; for these experiments, we use a value of 0.75, which we call the input firing mean (not 1 to allow the model to differentiate between x and x^2 in the plasticity rule). There is Gaussian noise with zero mean and a variance of 0.05 added to account for biological variability in the signal. In one trajectory, traditionally there are 240 trials, consisting of 3 blocks with different reward contingencies for odors in each block. The reward ratios are:

	Odor A	Odor B
Block 1	0.2	0.8
Block 2	0.9	0.1
Block 3	0.2	0.8

Updates are performed on a single trajectory, with no batching. In our simulations, we use 18 trajectories for training (matching the size of our previous experimental data) and 7 for evaluation for each seed. For each seed, results are reported as the median. Unless stated otherwise, all reported results are averaged over 3 seeds.

A.5 Additional experimental parameters

In the following subsections we explore the effect of three factors (regularization, moving average window, input firing mean) on model performance. Only the parameter being tested is varied while the other parameters are held fixed at the previously defined values.

A.5.1 L1 regularization

We experiment with various values of the L1 regularization penalty applied to the Taylor coefficients. This encourages sparse plasticity solutions and prevents the coefficients from exploding into NaNs due to the learning of positive values that exponentially increase the synaptic weight as the number of time points grows. We do not apply L1 regularization to the MLP parameters. The results of these experiments supports our original choice of L1 regularization level (Fig. 5).

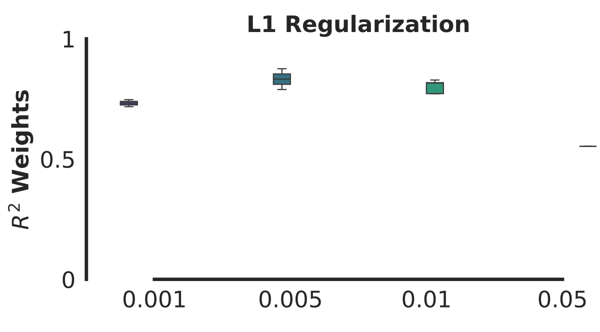

Figure 1: Effect of L1 regularization on R^2 of weights for Taylor plasticity rule

A.5.2 Moving average window

The moving average window refers to the window size used for calculating the expected reward. For example, a moving average window of 10 would take the average of the rewards received over the last 10 trials. Our exploration of window size suggests that smaller windows allow our model to more accurately predict weights in behavioral simulation experiments (Fig. 6). This is to be expected as shorter historical dependencies on past trials reduce the noisiness of the expected reward estimates. Importantly however, our model still robustly identified learning rules when using out default choice of 10 trials which was guided by experimental results from [**?**,]]rajagopalan2023reward that suggested that choice on a given trial was mediated by 10 past trials.

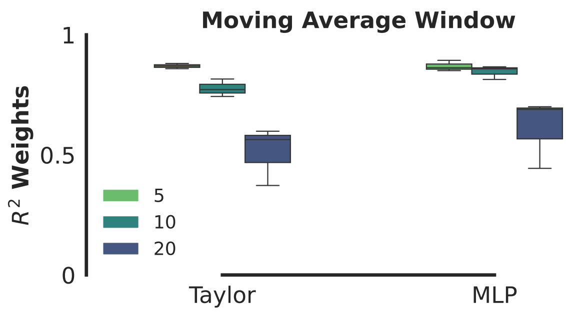

Figure 2: Effect of moving average window (used for calculated expected reward) on the performance of learned plasticity rule

A.5.3 Input firing mean

This is the firing mean encoding used for representing an odor. For example, a firing mean of 1 represents the odor as [1, 0], with additional Gaussian white noise fixed at a variance of 0.05. Our model (specifically the MLP implementation) is able to account for a range of input firing means (Fig. 7).

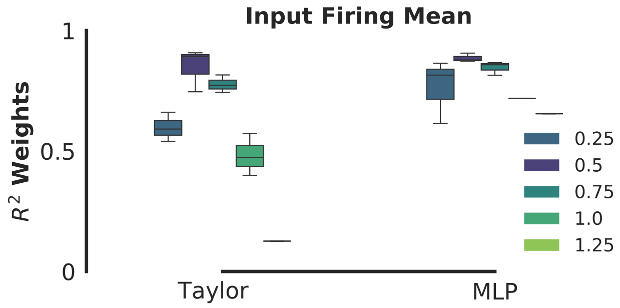

Figure 3: Effect of input firing mean (used for odor representation) on the performance of learned plasticity rule

A.6 Validation on held-out data

We ran a validation experiment to assess the extent to which the fitting procedure could generalize to unseen biological data. Due to the sequential nature of the dataset, and that we are fitting our model to individual flies' behavioral trajectories (*i.e.*, sequences of choices), we cannot perform classical k-fold cross-validation in which a random subset of timepoints or trials are held out. Instead, we train the model using the first $x\%$ of an individual fly's trajectory, and then we test it on the last $(100 - x)\%$ (**??**). To ensure there is no data leakage from the training set, we re-initialize the model's synaptic weights at the beginning of each test sequence, although the test performance is similar if the initial weights are carried over from the final timestep of the training sequence.

Interestingly, we found that some flies had test performance close to training performance, whereas others had poor test performance (**??**). The flies with good test performance (defined as having a positive percent deviance explained on the test set) reaffirmed our conclusion that a weight decay term enabled a quantitatively better description of fly behavior (**??**, top). The flies with poor test performance (negative test percent deviance explained) also had negative plasticity coefficients for the w term (θ_{001} in **??**), suggesting that the differences in test performance were not due to different estimates of this coefficient (**??**, bottom). More analyses are required to determine why some flies had markedly better test performance than others.

A.7 Additional plasticity rules

In these additional experiments, we maintain the reward term as the difference from the expected reward. This approach facilitates bidirectional plasticity. Additionally, we incorporate a weight decay term, experimenting with several

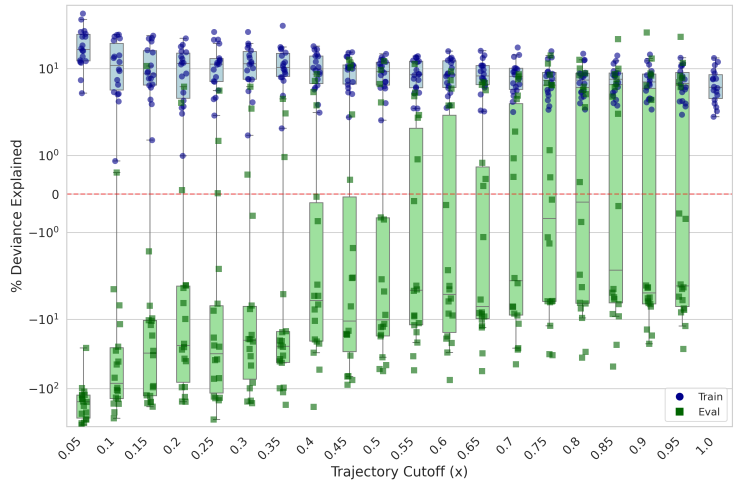

Figure 4: Percent deviance explained on the training and test data, training on $x\%$ of the fly data trajectory and testing on the remaining $100 - x\%$.

Figure 5: Learned weight decay term in the Taylor series parametrized plasticity rule (**??**), from flies with positive (top) and negative (bottom) test set percent deviance explained.

coefficients, ultimately choosing a value of 0.05 as it seems reasonable for our experimental configuration. **??** presents a comparison between the MLP and the Taylor series, reporting the R^2 over weights, R^2 over activity, and the percent deviance explained. At a high level, both methods appear to perform similarly. To gain deeper insights into why certain rules are more "recoverable" than others, an examination of weight dynamics for each method is necessary. The Taylor series model has 81 trainable parameters, while the MLP has 61. Fitting the rules with a relevant subset of the Taylor series, selected through biological priors as done in the *Drosophila* experimental data, is expected to result in better performance.

tableEvaluation of various different reward-based plasticity rules: R^2 scores for weight and individual neural activity trajectories, and percentage of deviance explained for behavior.

Plasticity Rule	MLP			Taylor		
	R^2 Weights	R^2 Activity	% Deviance	R^2 Weights	R^2 Activity	% Deviance
$x_j y_i w_{ij} r - 0.05r$	0.27	0.34	79.92	0.41	0.51	84.22
$x_j w_{ij} y_i r - 0.05r$	0.61	0.67	84.39	0.33	0.30	82.79
$x_j w_{ij} r^2 - 0.05r$	0.66	0.66	84.75	0.42	0.47	79.33
$x_j y_i w_{ij} r^2 - 0.05r$	0.40	0.57	80.56	0.40	0.48	79.03
$x_j r^2 - 0.05r$	0.75	0.73	70.78	0.66	0.61	70.65
$x_j^2 y_i w_{ij} r - 0.05$	0.59	0.68	70.41	0.59	0.69	68.84
$x_j^2 y_i^2 r^2 - 0.05$	0.59	0.62	64.98	0.55	0.62	64.44
$x_j r - 0.05 x_j y_i r$	0.81	0.96	63.64	0.81	0.95	64.01
$x_j r - 0.05 x_j w_{ij} r$	0.84	0.95	64.19	0.82	0.95	63.24
$x_j r - 0.05 x_j y_i$	0.80	0.96	62.90	0.76	0.94	62.50
$x_j r$	0.85	0.96	64.76	0.78	0.94	61.91
$x_j r - 0.05r$	0.70	0.76	62.01	0.63	0.68	61.82
$x_j r - 0.05 x_j y_i w_{ij} r$	0.82	0.96	63.13	0.73	0.93	61.05
$x_j r - 0.05 x_j w_{ij}$	0.83	0.95	61.24	0.81	0.95	60.85
$x_j r - 0.05 x_j y_i w_{ij}$	0.84	0.96	63.53	0.72	0.92	60.33
$x_j r - 0.05 x_j$	0.89	0.97	61.99	0.77	0.91	60.05
$x_j r^2 - 0.05 x_j w_{ij}$	0.89	0.95	54.92	0.88	0.94	53.94
$x_j r^2 - 0.05 x_j y_i$	0.79	0.95	56.40	0.77	0.92	53.76
$x_j r^2$	0.89	0.96	54.32	0.81	0.92	52.87
$x_j r - 0.05 w_{ij}$	0.87	0.91	57.01	0.70	0.86	51.80
$x_j r^2 - 0.05 x_j y_i r$	0.84	0.96	53.14	0.78	0.93	51.61
$x_j r^2 - 0.05 x_j y_i w_{ij}$	0.81	0.95	54.92	0.76	0.93	51.52
$x_j r^2 - 0.05 x_j w_{ij} r$	0.85	0.96	53.04	0.78	0.92	51.30
$x_j r^2 - 0.05 x_j y_i w_{ij} r$	0.90	0.96	53.90	0.82	0.92	51.29
$x_j r^2 - 0.05 x_j$	0.89	0.95	51.52	0.85	0.93	51.03
$x_j r^2 - 0.05 x_j r$	0.88	0.96	54.17	0.79	0.93	50.90
$x_j r - 0.05 y_i w_{ij} r$	0.91	0.95	52.94	0.84	0.91	50.24
$x_j r - 0.05 y_i r$	0.94	0.95	52.84	0.86	0.92	50.20
$x_j r - 0.05 w_{ij} r$	0.87	0.94	49.72	0.82	0.92	48.90
$x_j r - 0.05 y_i w_{ij}$	0.94	0.96	48.58	0.90	0.95	48.60
$y_i w_{ij} r^2 - 0.05$	0.66	0.90	47.26	0.55	0.84	47.00
$x_j r^2 - 0.05 w_{ij}$	0.86	0.94	46.44	0.78	0.91	46.26
$x_j r^2 - 0.05 y_i w_{ij} r$	0.91	0.92	43.64	0.89	0.93	44.04
$x_j r^2 - 0.05 y_i r$	0.96	0.97	43.74	0.93	0.95	43.18
$x_j r^2 - 0.05 w_{ij}$	0.92	0.93	42.68	0.91	0.94	42.43
$x_j^2 y_i w_{ij} r^2 - 0.05r$	0.65	0.53	43.26	0.65	0.57	42.32
$x_j r^2 - 0.05 y_i r$	0.93	0.92	44.13	0.91	0.94	41.69
$y_i w_{ij} r - 0.05$	0.69	0.89	38.52	0.61	0.85	37.74
$y_i^2 r^2 - 0.05$	0.52	0.87	37.01	0.60	0.82	36.81
$x_j r^2 - 0.05 y_i$	0.97	0.97	34.55	0.96	0.97	34.36
$x_j^2 y_i^2 r^2 - 0.05r$	0.72	0.65	33.94	0.68	0.60	33.57
$x_j^2 y_i w_{ij} r - 0.05r$	0.60	0.53	34.82	0.65	0.57	33.45
$y_i w_{ij} r - 0.05 x_j r$	0.04	0.45	28.56	0.08	0.46	28.63
$x_j^2 y_i^2 r - 0.05r$	0.50	0.44	28.81	0.70	0.60	28.08
$y_i w_{ij} r^2 - 0.05 x_j r$	0.17	0.47	26.90	0.31	0.44	24.92
$y_i w_{ij} r^2 - 0.05r$	0.38	0.32	16.49	0.60	0.59	15.87
$y_i^2 r^2 - 0.05r$	0.01	-0.01	14.06	0.24	0.29	13.62
$x_j^2 y_i^2 w_{ij} r - 0.05 x_j r$	0.82	0.94	4.64	0.86	0.95	4.35
$x_j^2 y_i^2 w_{ij} r^2 - 0.05 x_j r$	0.78	0.94	4.31 8	0.90	0.96	4.01
$x_j r - 0.05 x_j r$	0.75	0.92	4.30	0.92	0.96	3.96
$x_j y_i^2 w_{ij} r^2 - 0.05 x_j r$	0.79	0.93	4.39	0.89	0.97	3.87
$x_j y_i w_{ij} r - 0.05 x_j r$	0.70	0.91	3.49	0.87	0.96	3.29
$x_j y_i^2 r - 0.05 x_j r$	0.63	0.90	3.28	0.83	0.96	3.04

B NeurIPS Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

3. **Theory Assumptions and Proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [NA]

4. **Experimental Result Reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

6. **Experimental Setting/Details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

   Answer: [Yes]

7. **Experiment Statistical Significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [No]

Justification: All experiments are reproducible, and code is opensource. We report the rough time it takes to run an experiments, but have not provided extensive indepth details like the time of execution, etc.

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

10. **Broader Impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [NA]

    Justification: We have discussed the impact that our paper could have within the field of neuroscience. Larger societal impact of this work however is unclear, and likely will only be felt down the road as this approach is used more widely, we have therefore not dedicated a section to discuss this.

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

13. **New Assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

14. **Crowdsourcing and Research with Human Subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]