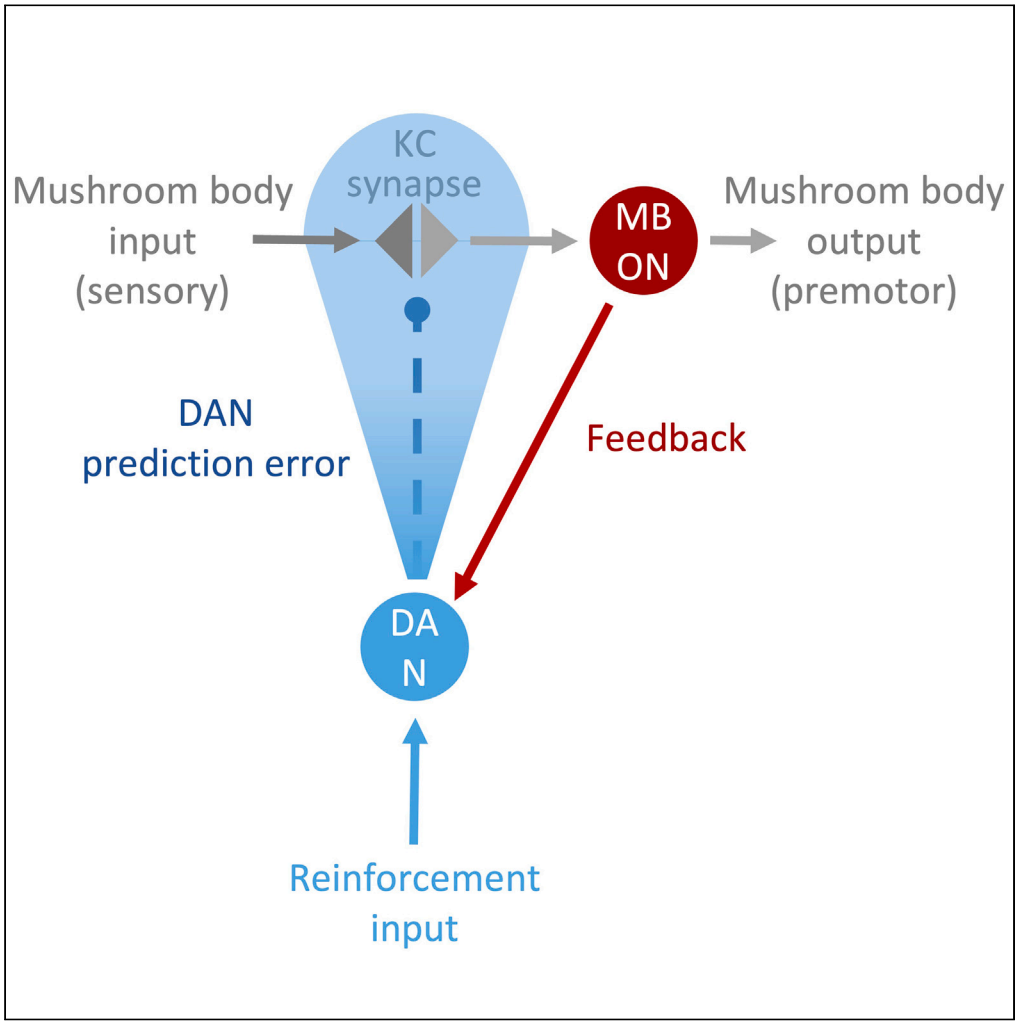
iScience

CellPress
OPEN ACCESS

Article

Prediction error drives associative learning and conditioned behavior in a spiking model of Drosophila larva



Anna-Maria Jürgensen, Panagiotis Sakagiannis, Michael Schleyer, Bertram Gerber, Martin Paul Nawrot

martin.nawrot@uni-koeln.de

Highlights

Dopaminergic neurons in a mushroom body model encode prediction error

Computation of prediction error is based on a feedback motif within the mushroom body

Prediction error can explain temporal dynamics of the acquisition of associations

Prediction error-driven learning can explain insect behavior

Jürgensen et al., iScience 27, 108640
January 19, 2024
https://doi.org/10.1016/j.isci.2023.108640

iScience

CellPress
OPEN ACCESS

Article

Prediction error drives associative learning and conditioned behavior in a spiking model of Drosophila larva

Anna-Maria Jürgensen,[1] Panagiotis Sakagiannis,[1] Michael Schleyer,[2,3] Bertram Gerber,[2,4,5] and Martin Paul Nawrot[1,6,*]

SUMMARY

Predicting reinforcement from sensory cues is beneficial for goal-directed behavior. In insect brains, underlying associations between cues and reinforcement, encoded by dopaminergic neurons, are formed in the mushroom body. We propose a spiking model of the *Drosophila* larva mushroom body. It includes a feedback motif conveying learned reinforcement expectation to dopaminergic neurons, which can compute prediction error as the difference between expected and present reinforcement. We demonstrate that this can serve as a driving force in learning. When combined with synaptic homeostasis, our model accounts for theoretically derived features of acquisition and loss of associations that depend on the intensity of the reinforcement and its temporal proximity to the cue. From modeling olfactory learning over the time course of behavioral experiments and simulating the locomotion of individual larvae toward or away from odor sources in a virtual environment, we conclude that learning driven by prediction errors can explain larval behavior.

INTRODUCTION

Associative learning is a fundamental cognitive ability across vertebrate[1–4] and invertebrate species.[5–8] In insects, the mushroom body (MB) is a central brain structure for multi-sensory integration, involved in memory formation and recall.[9,10] The synapses between its intrinsic and output neurons are at the core of memory formation during associative learning.[6,9,11–14]

During associative learning, a relationship between two previously unrelated elements is established gradually. In classical conditioning, representing one form of associative learning, a so-called conditioned stimulus (CS), obtains behavioral relevance through its concurrence with the reinforcing unconditioned stimulus (US). The temporal evolution of this memory acquisition process depends dynamically on the spatiotemporal proximity of the CS and US, as formalized in the Rescorla-Wagner (RW) model.[15]

$$\begin{aligned} \Delta V &= \alpha \cdot (\lambda_{US} - V(t)), \\ V(t+\Delta t) &= V(t) + \Delta V. \end{aligned} \qquad \text{(Equation 1)}$$

The asymptote λ_{US} of the learning curve defines the point at which the CS fully predicts the US and the acquisition of their association terminates. Until then, the change in associative strength ΔV is proportional to the difference between the maximum associative strength λ_{US} and the already acquired associative strength $V(t)$. This driving force in the learning process has been termed prediction error (PE)$PE = \lambda_{US} - V(t)$.[16] Throughout the memory acquisition phase, the pace of learning, which can be formalized as the slope of the acquisition curve, decreases as the PE becomes smaller, reducing the driving force for changes of the association $V(t)$.[15–17] This continuous update of reinforcement predictions, guided by the PE, could allow animals to adapt their anticipatory behavior to the predicted US efficiently[18,19] and thus approach reward or avoid punishment.

Dopaminergic neurons (DANs) have long been known to encode information about reward and punishment. These types of neurons respond to the presence of rewards and punishments in the environment, shown both in vertebrates[20–25] and *Drosophila*.[26–29] The activation of DANs induces approach or avoidance learning as has been found in vertebrates[30–33] and larval[28,34–36] and adult[37–39] *Drosophila*. This approach or avoidance learning is facilitated by the modulation of the synapses between intrinsic and output neurons of the MB.[11,13,14,40] Ultimately DANs do not only signal the presence of rewards or punishments but have also been suggested to encode a PE in various vertebrate species[24,41–44] and might have a similar function in insects.[27,28,40,45–50]

[1]Computational Systems Neuroscience, Institute of Zoology, University of Cologne, 50674 Cologne, Germany
[2]Leibniz Institute for Neurobiology (LIN), Department of Genetics, 39118 Magdeburg, Germany
[3]Institute for the Advancement of Higher Education, Faculty of Science, Hokkaido University, Sapporo 060-08080, Japan
[4]Institute for Biology, Otto-von-Guericke University, 39120 Magdeburg, Germany
[5]Center for Brain and Behavioral Sciences (CBBS), Otto-von-Guericke University, 39118 Magdeburg, Germany
[6]Lead contact
*Correspondence: martin.nawrot@uni-koeln.de
https://doi.org/10.1016/j.isci.2023.108640

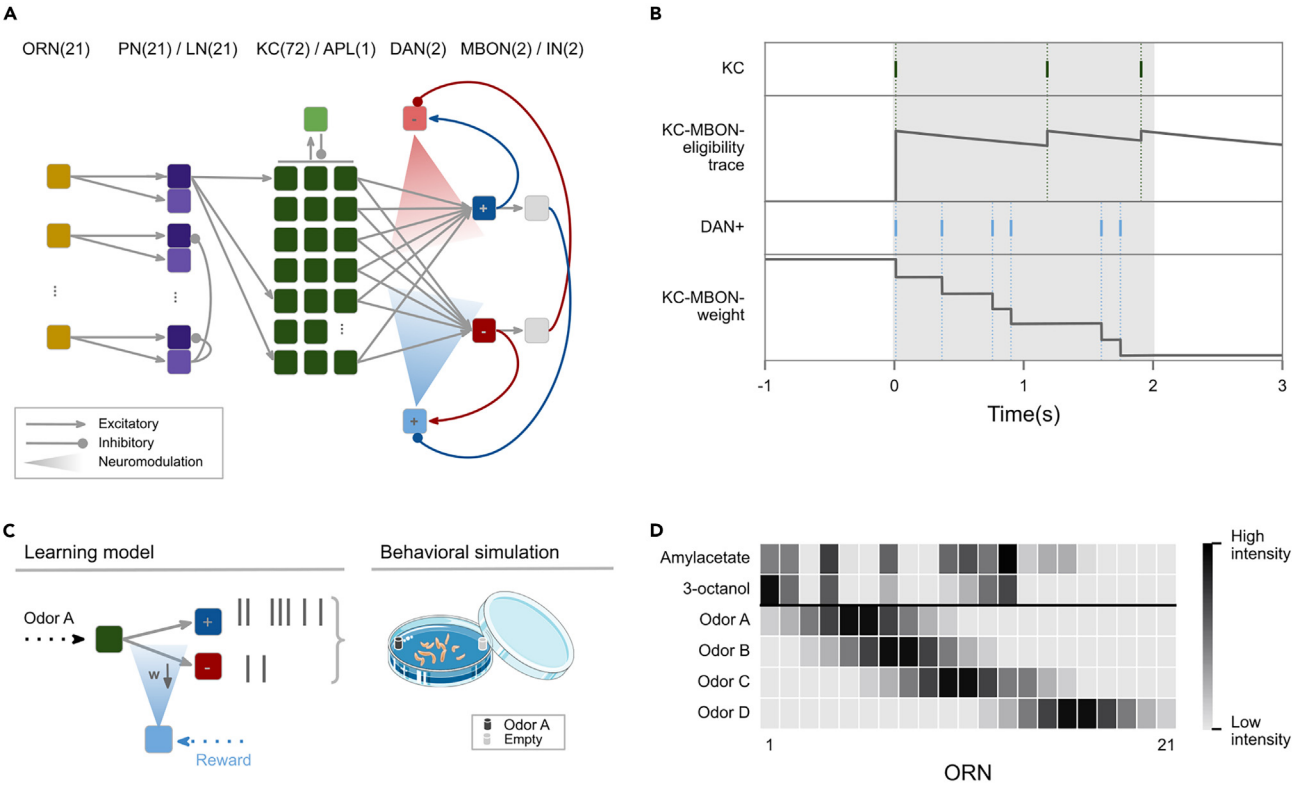


Figure 1. Network mechanisms

(A) Network model of the *Drosophila* larva olfactory pathway including all neurons and connections implemented. One-to-one feedforward connections between 21 olfactory receptor neurons (ORNs) and 21 projection neurons (PNs)/local interneurons (LNs) and from 2 to 6 PN to each of the 72 Kenyon cells (KCs). Lateral inhibition from each LN innervates all PNs, and recurrent feedback inhibition from the anterior paired lateral (APL) neuron is provided onto all KCs. The MB output region is organized into two distinct compartments. The upper compartment holds the approach encoding $MBON_+$ and is innervated by the punishment-mediating DAN_-; the lower compartment holds the avoidance mediating $MBON_-$ and is innervated by the reward-mediating DAN_+. Each DAN can exert a neuromodulatory effect on the plastic KC>MBON synapses within its compartment. MBONs provide excitatory and inhibitory (via gray interneurons, LN) feedback to the DANs. The implementation of the olfactory pathway with ORNs, PNs, LNs, APL, and KCs is similar to a previous model that lacks plasticity and the MB output circuitry.[1]

(B) Sketch of synaptic weight change at a single KC>MBON synapse with respect to the synaptic eligibility trace elicited by KC spikes and the occurrence of reward-triggered spikes in DAN_+. Amylacetate is paired with a reward for 2 s (gray shaded area).

(C) To generate simulated larval behavior in the Petri dish during the test phase of the learning experiments, we utilized our locomotory model,[2] based on the behavioral bias (see method details in STAR Methods) acquired by the MB model during the training phase. The behavioral bias is used directly as input to the locomotory model.

(D) All odors (see method details in STAR Methods) were used in the experiments: naturalistic odor patterns for amylacetate and 3-octanol, as well as four artificial patterns (odorA, odorB, odorC, odorD) with varying distances (see method details in STAR Methods) from odorA. Each odor activates a different set of input neurons with a different spike rate, as indicated by the scale bar.

We introduce a biologically realistic spiking model of the *Drosophila* larva olfactory pathway and MB in one brain hemisphere that forms associations of odors with reward to further test the hypothesis that within this circuit PE coding takes place in DANs that receive feedback from the output neurons of the MB or their downstream partners through either direct or indirect pathways.[12,28,51–54] Beyond the scope of similar models[28,55–58] (see discussion), we demonstrate that the proposed mechanism can reproduce the experimentally observed findings on the acquisition of associations of odors with reward in a time-resolved manner.[59] To facilitate direct qualitative and quantitative comparisons with animal behavior, we couple our MB model with a realistic locomotory model of the larva[60] that captures the learning-induced adaptation of chemotactic behavior in the individual animal.

RESULTS

Connectome-based circuit model of the larval olfactory pathway

The network architecture of our model (Figure 1A) connects our previous olfactory pathway model[61] with a plastic circuit model of the MB and its output neurons. Our model is based on the anatomy of the olfactory pathway and the MB in a single *Drosophila* larva brain hemisphere[28,34,62,63] (see method details in STAR Methods). Peripheral processing is carried out by 21 olfactory receptor neurons (ORNs), each

expressing a different olfactory receptor type.[62,64,65] ORNs form one-to-one excitatory synaptic connections with 21 projection neurons (PNs) and 21 local interneurons (LNs) in the antennal lobe.[62] Each LN connects with all PNs via inhibitory synapses, establishing a motif for lateral inhibition within the antennal lobe. The 72 mature larval Kenyon cells (KCs)[63] are the excitatory intrinsic neurons of the MB. Each KC receives excitatory input from 2 to 6 randomly selected PNs. Additionally, a few KCs receive input from a single PN.[63] These one-to-one PN>KC connections are excluded from our circuit model. The KCs are subjected to feedback inhibition, provided via the GABAergic anterior paired lateral (APL) neuron, which receives input from all KCs.[34] Only mature KCs, characterized by a fully developed dendrite, are included in this model, yielding a complete convergent synaptic KC>APL connectivity. The MB lobes are organized in compartments, in which the KC axons converge with the dendrites of one or few MB output neurons (MBONs).[28,63] Our model assumes two MBONs from two different compartments that are representative of two different categories of output neurons of the MB that mediate either approach or avoidance (Figure 1A).[11–14,66–69] Both MBONs receive excitatory input from all of the KCs to fully capture the information that is normally represented by the complete set of MBONs. Each compartment is also innervated by a single DAN, signaling either reward (DAN_+) or punishment (DAN_-) that targets the KC>MBON synapses to facilitate learning.

Learning and synaptic plasticity at the KC>MBON synapses

We assume that the KC>MBON synapses undergo plasticity (Figures 1A and 1B), based on strong experimental evidence in larval[28,66,67] and adult flies.[11,13,14,40] This plasticity requires the convergence of the sensory pathways in the form of KC activation and of the reinforcing pathway, mediated by a neuromodulatory DAN signal at the synaptic site. We employ the two-factor learning rule

$$\Delta w_i^l = -a \cdot e_i(t) \cdot R(t) \leq 0 \qquad \text{(Equation 2)}$$

to model learning-induced plasticity at each KC>MBON synapse i (Figures 1A and 1B). The first factor is expressed in the presynaptic KC activation by an odor, rendering the synapse eligible for modification. This is modeled via an exponentially decaying eligibility trace $e_i(t)$, which is set to one whenever the respective KC elicits an action potential (Figure 1B). Its decay time constant determines the window of opportunity for synaptic change. The presence of a reinforcing stimulus constitutes the second factor and is signaled by the reward-mediating DAN_+ or punishment-mediating DAN-. Spiking of the DAN_+ (or DAN_-) provides a neuromodulatory reward (or punishment) signal $R(t)$ to the synaptic site. When a DAN action potential coincides with positive eligibility at a given synapse i, the respective synaptic weight is reduced (Figure 1B) where the reduction Δw_i^l is proportional to the current value of the eligibility and is additionally controlled by the learning rate $a > 0$ (Table S1). A higher or lower DAN activity thus results in a more or less frequent reduction of the synaptic weight elicited by DAN spikes. Each synaptic weight is restricted to non-negative values.

Synaptic homeostasis

To account for the experimentally observed decline of a learned association when the reward is omitted[45,70–73] and to maintain memory capacity across the population of KC>MBON synapses, we introduce a mechanism for homeostatic synaptic plasticity[74,75] as

$$\Delta w_i^h = (w_{init} - w_i(t)) \cdot M(t) \cdot h \geq 0. \qquad \text{(Equation 3)}$$

This rule has two features. First, the magnitude of the homeostatic effect depends on the postsynaptic spiking activity $M(t)$ of the MBON. Second, we assume that the homeostatic regulation acts locally at individual synapses.[76,77] Each MBON action potential leads to a homeostatic increase $\Delta w_i^h > 0$ of the individual KC>MBON synaptic weight if $w_i(t) < w_i^h$. This increase is proportional to the extent to which the weight differs from its original value w_{init} (Table S1) and thus balances the learning-induced reduction of the synaptic weight. The homeostatic regulation factor $h \geq 0$ (Table S1) parameterizes the magnitude of the homeostatic effect. This mechanism serves as an implementation for the loss of the association when the reward is omitted, partly counteracting the plasticity effect at each KC>MBON synapse (see discussion, section "a mechanistic implementation of the RW model"), and also maintains memory capacity by counteracting the accumulated reduction of input weights over the course of the learning process (Equation 2). The sum of the learning-induced synaptic plasticity and the homeostatic regulation defines the magnitude of the synaptic weight at the next simulation time step $t + \Delta t$ as

$$w_i(t + \Delta t) = w_i(t) + \Delta w_i^l + \Delta w_i^h \qquad \text{(Equation 4)}$$

and thus the dynamic weight changes throughout an experiment.

Behavioral bias at the MB output

It has been shown experimentally that specific MBONs encode a behavioral tendency to either approach or avoid a currently perceived stimulus, depending on the acquired stimulus valence in adult[12–14,68,69] and larval[66] *Drosophila*. In the naive state of our model, all KC>MBON synapses have the same initial weight w_{init} (Table S1), and hence the spiking activity of both MBONs is highly similar. Learning alters the KC>MBON synaptic weights and thus skews the previously balanced MBON output. This acquired imbalance between MBON outputs biases behavior toward the approach or avoidance of the conditioned odor (Figure 1C). To quantify the effect of learning, we compute the behavioral bias toward odors used in the experiments from the spike count of both MBONs over a time interval T (also see method details in STAR Methods)

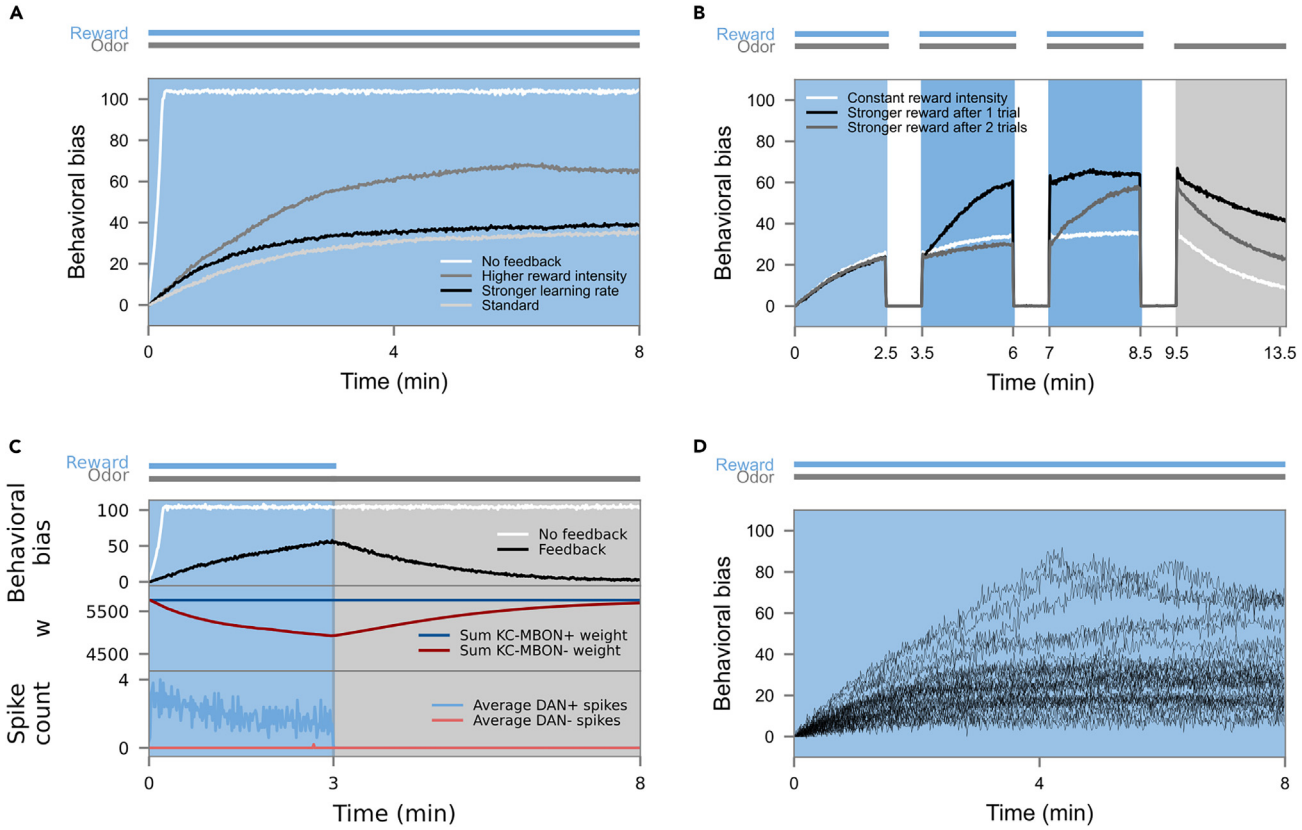


Figure 2. Learning with prediction errors

(A) $N = 30$ model instances were trained with the odor amylacetate (CS) and reward (US, blue background). MBON>DAN feedback, the reward/odor intensity, and the learning rate were manipulated in separate experiments. The odor preference (behavioral bias, see method details in STAR Methods) was measured continuously in windows of 1 s and averaged over all model instances.

(B) $N = 30$ model instances were trained during three trials with amylacetate and reward (blue background). Reward intensity was either constant across the three training trials (white curve) or enhanced during the third (gray) or the second and third trials (black). The training was followed by a 3 min test phase with odor only (gray background).

(C) $N = 30$ model instances were trained with amylacetate and reward (blue background) and then underwent an extended test phase (gray background).

(D) Individual acquisition curves for $N = 30$ model instances (standard experiment Figure 2A).

$$BB = \frac{MBON_{+} - MBON_{-}}{T}. \qquad \text{(Equation 5)}$$

Implementation of prediction error coding in the KC-MBON-DAN motif

In the larva, DANs and other modulatory neurons can receive excitatory and inhibitory input from different MBONs, either directly or via a single or two interneuron steps.[28] Based on this anatomical evidence, we propose a basic MBON>DAN feedback motif in Figure 1A (for similar models, see discussion section: Comparison with other MB models). In our model, DANs are activated by external reward/punishment signals and also receive direct excitatory (red) and indirect inhibitory (blue) feedback from both MBONs. As the initial balance between the two MBON outputs shifts over the course of the training process, the amount of excitatory and inhibitory feedback that DANs receive continues to diverge, allowing the DANs to access the model's learning history. Ultimately, DAN activation signals the PE as the difference between the current external activation by reinforcement in the situation and the expected reinforcement based on prior learning. Prior learning is implemented as the difference between excitatory and inhibitory MBON>DAN feedback.

The suggested feedback motif led to learning curves that saturated when the reward was fully predicted, and the PE approached zero (Figure 2A). This effect disappeared when the feedback circuit was disabled (Figure 2A). In this case, the behavioral bias quickly reached its maximum value when the MBON$_-$ elicited no more spikes. Increasing the reward intensity caused the learning curve to saturate at a higher level (Figure 2A) because $R(t)$ (Equation 2) was positive more often during each analysis time window of 1 s (Equation 5). Increasing the learning rate α by 65% fostered a faster acquisition of the association (Figure 2A). When changing the learning rate, we also adapted the homeostatic regulation factor h (Equation 3) by the same factor since it was fitted together with the learning rule. More information about the

isolated effect of synaptic homeostasis can be found in the supplemental information (Figure S1). Increasing the reward intensity after 2.5 min (black curve), or 5 min (gray curve), of appetitive training resulted in a steeper slope of the learning curve. Additionally, it increased the maximum during training trials of 2.5 min duration with increased reward intensity (Figure 2B). Higher intensity of the reward resulted in an average DAN spike rate of 39.14 ($\pm$1.27 SD) Hz compared to 33.11 ($\pm$1.34) Hz.

We next tested for the loss of the acquired association throughout prolonged exposure to the CS without the US following initial memory acquisition, as observed experimentally.[15,78] To test this in our model experiments, we now presented the odor that was previously paired with reward in the absence of reward and for an extended period. The extinction mechanism was no longer outweighed by learning and drove each individual weight back toward w_{init} (Figure 2C). We also demonstrated the interaction of the learning rule with this mechanism in Figure S1, where the learning rate remains constant. At the same time, the magnitude of the homeostatic regulation was manipulated to show that both mechanisms need to be in balance.

The observed properties of the learning curves are not just the result of averaging across many model instances but can be attributed to the saturating effect of the MBON>DAN feedback on each individual learning curve (Figure 2D).

Learned preference and behavior generalize to similar odors

We trained our model by pairing a reward with a single odor for 4 min. After the training procedure, we tested the behavioral bias for either the same or a different odor, following the experimental approach used in the larva.[79] Mimicking the experimental data, we showed that the odor preference is highest if training odor and test odor were identical in the case of training with 3-octanol. When amylacetate was used during training, 3-octanol preference was increased (Figure 3A). Figure 3B shows the network response to 30 s stimulations with amylacetate and 3-octanol in a single exemplary model instance. On the level of the ORNs, 3-octanol merely activated a subset of the amylacetate-activated neurons, some with a higher rate than in the case of amylacetate (Figure 1D). The uniqueness of the odor response pattern is enhanced in the KC population.[61] We systematically tested for generalization using a set of artificially designed ORN activation patterns with a controlled degree of overlap (see method details in STAR Methods, Figure 1D) and found that, with decreasing similarity, the generalization effect to a new odor was diminished (Figure 3A).

The model reproduces temporal features in trace conditioning experiments

Including an odor-evoked eligibility trace at the KC>MBON synapses allowed the model to maintain the sensory odor representation for a time window during which reward would trigger synaptic change (Figure 1B). The time window between odor and reward onset (0, 10, 20, 30, 40, 50, 60, 120 s) was varied for trace conditioning experiments with a 30 s presentation of odor and reward that was repeated three times. A small inter-stimulus interval (ISI) between the onset of the odor stimulus and the onset of the reward stimulus of 10 to 30 s led to an increase in behavioral bias compared to the complete overlap of odor and reward (Figure 3C), using the extended window of opportunity for synaptic change triggered by each KC spike. Long ISIs did not lead to learning as the eligibility trace had declined close to zero (Figure 3C). These findings match observations from experiments in larvae[34,80,81] with the caveat that the trace in the real larva brain seems to extend for a slightly longer period of time compared with our experiments.

The model reproduces paired and unpaired associative conditioning experiments

To test if learning, driven by PE, can account for learned larval behavior, we replicated a set of single-trial conditioning experiments performed with larvae[59] in simulation. In these experiments, simulated animals were first trained with the odor amylacetate in a single trial of varying duration. In the animal experiments, larvae are placed on a Petri dish coated with an agar-sugar substrate and the odor in two small containers for diffusion in the air (paired training, Figure 4A). Either before or following this training protocol, larvae undergo a single trial without sugar and odor. Afterward, the animals are transferred to a new dish with two odor containers placed on different sides. One of them contains amylacetate, and the other one is empty. This paired training was compared with an unpaired protocol (Figure 4A) with separate (randomized order) presentations of amylacetate and sugar. After the paired training protocol, the animals tend to approach the previously rewarded odor, as measured by the difference in the number of animals on each side at the end of a 3 min test phase, divided by the total number of animals. Following the experimental literature, we will refer to this measure as the preference index[59] (Equation 17 in quantification and statistical analysis in STAR Methods).

We aimed to replicate these behavioral experiments on two levels. First, we focused on the direct model output that reflects the strength of the acquired association between amylacetate and reward (behavioral bias, Equation 5, see also quantification and statistical analysis in STAR Methods) and later also simulated behavior based on these biases. We simulated both the paired and unpaired training protocols (Figure 4B). While the unpaired training yielded almost no behavioral bias, the models that underwent the paired training showed an increased behavioral bias that depended on the duration of the training and saturation for longer training duration (Figure 4B). The simulation results reported in Figure 4B were obtained using odor-naive models that exhibited no odor preference prior to training. To account for the experimental finding that real larvae typically express an odor preference even without any training,[82–84] we readjusted our experiments to include a pre-training period of 10 min to start the conditioning experiments with the amylacetate-reward association already established. This adaptation of the protocol led to new results (Figure 4C). The paired condition in Figure 4C showed that, once the behavioral bias was saturated (Figure 2A), continued pairing maintains the association without further increasing it. Unpaired training, however, caused the behavioral bias to decrease and saturate at a lower level. For a discussion of different potential causes of a reward expectation prior to training, please refer to the discussion (comparison of modeling results to experimental findings). Figure 2A shows that disabling MBON>DAN feedback led to a

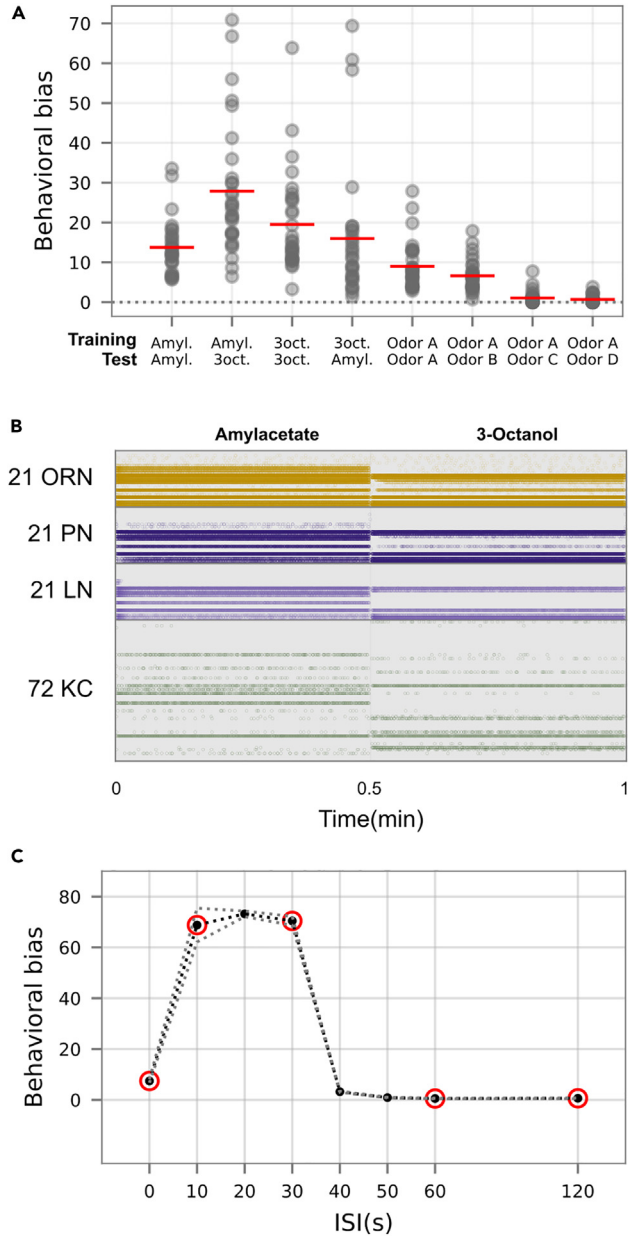


Figure 3. Reward generalization and trace conditioning

(A) The behavioral bias (see Method details in STAR Methods) generalizes to odors that differ from the training odor after a 4 min training (3 min test phase). We conducted simulation experiments with different combinations of training and testing odor, each for 10 groups (gray circles represent the mean of a single group) of $N = 30$ larvae, and red lines indicate the mean between groups. The behavioral bias is highest when the training and the testing odor are the same.

(B) Spiking activity in the network during the presentation of amylacetate (left) and 3-octanol (right) in a single naive model instance.

(C) Simulated trace conditioning experiments with odor (amylacetate) and reward. Inter-stimulus interval (ISI) indicates the time between odor and reward onset. The black line displays the mean, and gray lines the std over $N = 10$ groups of 30 model instances each. Conditions circled in red correspond to the conditions also used in animal experiments.[3,4]

learning curve that did not saturate but instead increased with a steep slope until it reached the maximum value for the behavioral bias (Equation 5) with a $MBON_{-}$ rate of 0. To verify if this PE feedback mechanism is responsible for the difference between maintenance and loss of the association in Figure 4C, we repeated the same experiment with disabled MBON>DAN feedback. The behavioral bias overall was much higher compared to the intact network (Figure 4D). The maximum was reached before the test phase of even the shortest experiment with 1 min training.

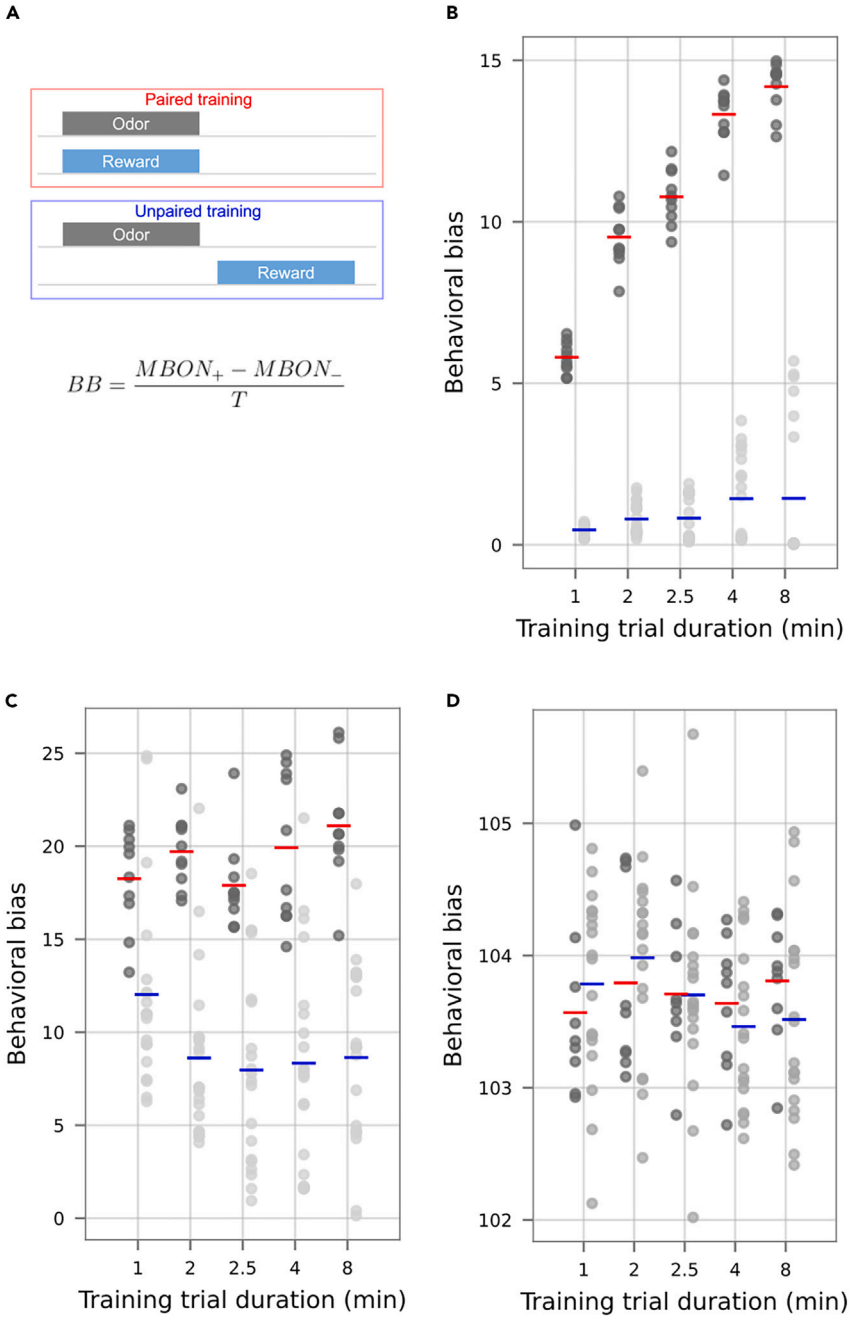


Figure 4. Paired and unpaired learning in the MB model

(A) Schematic overview of the paired vs. unpaired training protocol.

(B) The model's behavioral bias (see method details in STAR Methods) for training with amylacetate and reward for $N = 10$ paired (dark gray, mean in red) and $N = 10$ unpaired (light gray, mean in blue) experiments with groups of 30 modeled larvae each. In the unpaired condition, half of the groups were trained with the odor preceding the reward. For the other half, the reward preceded the odor.

(C) Model behavioral bias for amylacetate for $N = 30$ paired and $N = 30$ unpaired experiments with randomized order of odor and reward. Prior to the conditioning experiment, the model instances underwent a 10 min pre-training period, during which odor and reward were paired.

(D) Model behavioral bias for amylacetate for $N = 30$ paired and $N = 30$ unpaired experiments with randomized order of odor and reward. The MBON>DAN feedback was disabled. Before the conditioning experiment, the model instances underwent a 10 min pre-training.

After examining the readout of the MB directly (Figure 4), we next performed behavioral simulations of the testing phase with groups of virtual larvae for both the paired and unpaired conditions.[59] Since the effect of training in lab experiments was quantified behaviorally via spatially defined, group-level metrics (preference index[59] and performance index,[59]; Equations 17 and 18), this allows a straightforward comparison between the animal (Figures 5A and 5B) and simulation experiments (Figures 5C and 5D). To this end, we utilized a realistic model for the simulation of larval locomotion and chemotactic behavior[60] that uses the behavioral bias at the output of the MB model as a constant factor, throughout the test to modulate the locomotory behavior of individual larvae toward or away from a spatially placed odor source in a virtual arena (see quantification and statistical analysis in STAR Methods). The resulting preference indices, acquired across groups of independently simulated larvae (Figure 5C), can be directly compared to the experimentally obtained preference indices (Figure 5A). We also compared performance indices from our simulated experiments (Figure 5D) with those from the lab experiments (Figure 5B). We found that the model can replicate these when assuming an odor preference at the beginning of the experiment. The animals' preference is relatively consistent across training trials of different duration. Prolonged paired training did not lead to an increase in preference (Figure 5A). These experiments did not include a test for odor preference before training, but the naive larval preference of odors used in learning experiments has been demonstrated elsewhere repeatedly.[82–84] This paired training was compared with an unpaired protocol with separate (randomized order) presentations of amylacetate and sugar. Here, the extent to which animals preferred amylacetate over no odor varied with the duration of the training trial. The longer the duration of the training, the more the preference index decreased from an initially high value but saturated around 2.5 min (Figure 5A).

DISCUSSION

We proposed a basic PE encoding circuit motif in the larval MB and tested its biological plausibility and capacity to explain behavioral data in a continuous simulation using our spiking network model of the olfactory pathway[61] and MB. We demonstrate that this mechanistic model for PE coding results in saturating individual and group-level learning curves where the slope and maximum of the learning curve are determined by the intensity of the reward. Learning is also influenced by the relative timing of odor and reward and can be extinguished if the reward is omitted during the presentation of the sensory cue. Coupling of the spiking neural network simulation with the simulation of larval locomotory behavior allowed to explain behavioral results obtained in larval learning experiments.

A mechanistic implementation of the RW model

A number of predictions can be derived from the phenomenological RW model[15] and tested in our mechanistic model thereof. We found that, regardless of odor and reward intensities or the model's learning rate, the strength of the odor-reward association (quantified by the behavioral bias, Equation 5) saturated over time (Figure 2A), as the strength of the already acquired association $V(t)$ approached the maximum value determined by the present reward (λ_{US}) (Equation 1). Consequently, our model's acquisition curve saturates at a higher value when the intensity of the reward is increased (Figures 2A and 2B), as predicted by the RW model, in which a stronger US should result in a higher value of the asymptote λ_{US}.[15] In our model, a higher reward intensity relates to a higher input rate to the respective DAN (see method details in STAR Methods), which translates into more frequent DAN action potentials within a given window of 1 s. According to the RW model, increasing the learning rate α[15] should lead to faster acquisition of the association without changing the asymptote. In our model, the learning rate parameter enters the synaptic learning rule and directly modulates the increment of the synaptic weights (Equation 2).

The RW model predicts that the omission of reward should result in the loss of the learned association[15] (Equation 1). At this theoretical level, we cannot infer if this loss is due to extinction or forgetting. Extinction, characterized by the possibility of recovery of the association after its temporary loss,[85] has been demonstrated in adult[73,86] but not yet in larval *Drosophila* (for a discussion see Mancini et al.[87]). To retain the association for recovery, extinction relies on the formation of parallel memory traces for the maintenance and the loss of the association.[45,72,88] The mechanism implemented in our model (Equation 3) is a sensory input-dependent homeostatic mechanism of forgetting that slowly pushes all synaptic weights back toward their initial value. This steadily deletes any previously learned association without the possibility of recovery, but only in the presence of olfactory input. Note that, because the postsynaptic MBON activity $M(t)$ enters Equation 3, any olfactory input may trigger homeostatic forgetting. As a consequence, if the model is first trained by pairing odor A with reward and subsequently undergoes a second pairing of odor B with reward, then the reward association with odor A is diminished during the acquisition of the reward association with odor B, providing an experimentally testable prediction.

Comparison of modeling results to experimental findings

A variety of experiments have demonstrated group-level acquisition curves that saturate over multiple training trials or with increasing duration of a single trial in olfactory conditioning.[58,59,71,89,90] To replicate larval behavior in reward learning experiments[59] with varying duration of the learning phase (Figures 5A and 5B), we trained our model with an odor and reward in a paired vs. unpaired fashion (Figure 4B). Real larvae exhibit a naive (often positive) preference for many odors.[82,91–93] They demonstrate a strong odor preference after a very short training and no significant increase in their preference when trained in a paired manner for longer periods.[59,82] In contrast, animals trained in an unpaired protocol start with a similarly high odor preference, which then decreases over time.[59,82] This observation from animal experiments is somewhat counter-intuitive since the coincidence of odor and reward should yield an increase in the association of those two stimuli,[15] which can be observed in our modeling results (Figure 4B). To resolve this, we included the observation that odors are not always neutral to untrained animals and assumed odor preference before the conditioning experiment.[59] In this case, the animals would enter into the experiment with an

$$Pref = \frac{count_{odor} - count_{no\ odor}}{count_{odor} + count_{no\ odor}}.$$

$$PI = \frac{Pref_{paired} - Pref_{unpaired}}{2}.$$

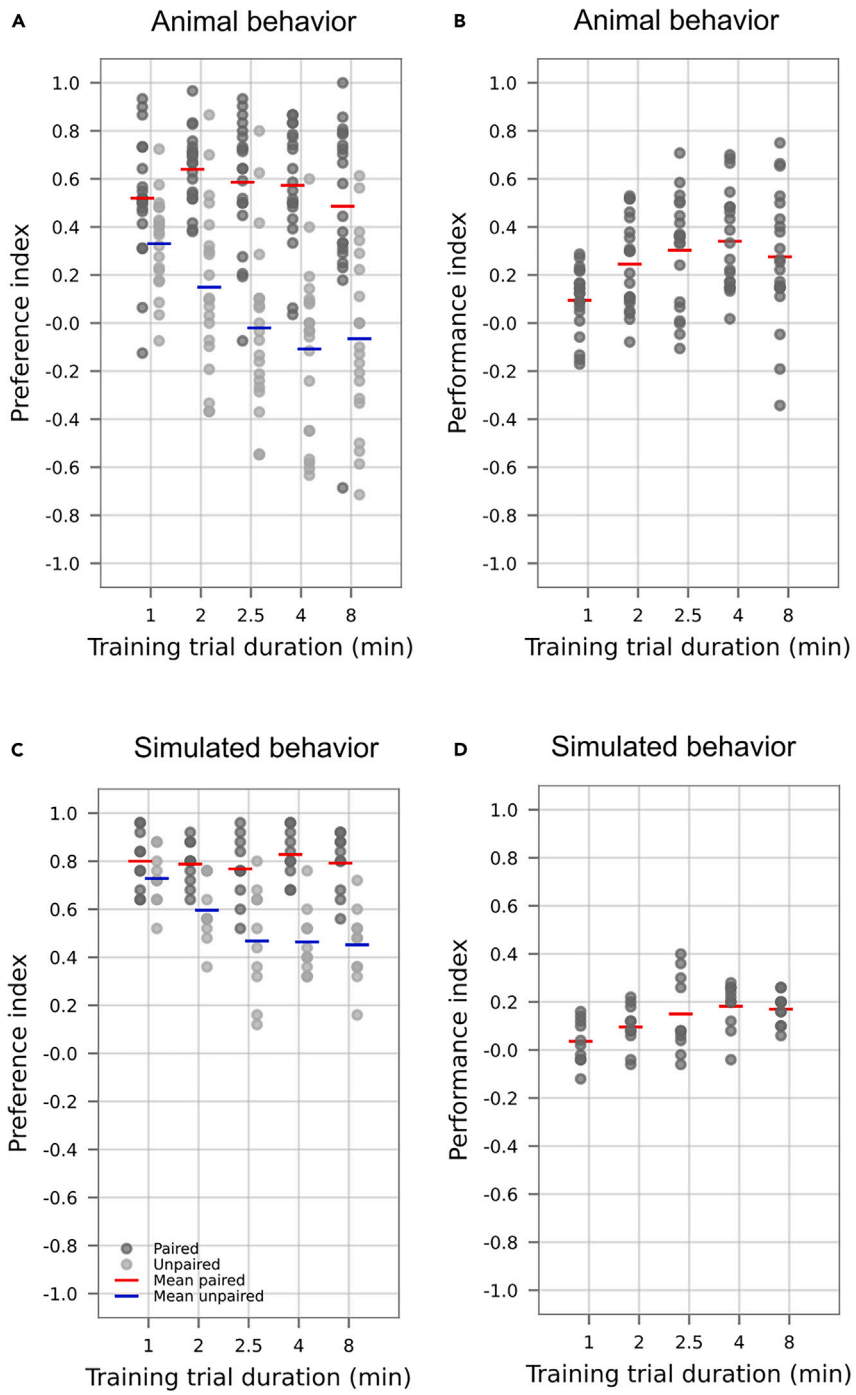


Figure 5. Replicating behavioral experiments with paired and unpaired training

(A) Experimental preference indices for amylacetate for 20 groups of 30 real animals each for paired and unpaired experiments with randomized order of odor and reward (data published by Weiglein et al. (2019),[5] https://learnmem.cshlp.org/content/26/4/109/suppl/DC1).

(B) Experimental performance indices for amylacetate computed between preference in paired and unpaired real animal experiments.[5]

(C) The simulated behavior is based on the protocol in A. Simulated preference indices for amylacetate for $N = 10$ paired and $N = 10$ unpaired experiments with varied order of odor and reward.

(D) Simulated performance indices for amylacetate computed between preference in paired and unpaired simulation experiments.

already established reward prediction that would be violated during unpaired training. Three scenarios lend themselves as plausible causes of this effect. First one is accidental conditioning over the course of their lifespan, during which they are raised on a food substrate while being exposed to air that carries many different odorants. Alternatively, or in fact, additionally, the animals likely exhibit innate preferences for several odors.[91,92,94] Finally, the presence of the reward during reward-only phases might lead to an association of the experimental context with that reward (previously discussed by Saumweber et al.[82]). In our model, this could be realized by providing context-specific multi-sensory input to the MB that can be associated with the model solution suggested by Müller et al.[95] The resulting reward expectation (solely based on the always-present context), unmet during the odor-only phases, could then lead to a negative PE signal. All three candidate explanations would yield a similar projection for the unpaired experimental protocol: a reward expectation acquired prior to the actual experiment would cause a violation of that expectation during odor-only trials of the unpaired experiments. In all three cases, the animal's preferences might also generalize to a broader array of odors, leading to an overall preference for some odors, as observed experimentally. To test this hypothesis, we pre-trained our model before simulating conditioning experiments (Figure 4C) to establish an odor preference at the beginning of the experiment. This not only ensured that the model behaves in accordance with the RW model[15] but also fits the animal experimental results[59] (Figures 5A and 5B). A possible alternative explanation could be a sensory habituation process to the odor that might cause odor preferences to decrease over time in the unpaired condition, while, in the paired condition, this effect might be counteracted by the continued presentation of odor and reward together.[96]

Thus far, we tested our model in experiments where the CS and US presentations fully overlapped (paired conditions). We now consider different onset times, with the onset of the CS always preceding the onset of the US (Figure 3C). For these experiments, we used a shorter duration of 30s for both CS and US presentation, repeated over three acquisition trials to mimic experimental conditions in larval experiments[34,80,81] that used optogenetic activation of DANs as a proxy for sugar reward (or punishment). Similar to their experiments, we show that the behavioral bias clearly depends on the temporal delay between CS and US (Figure 3C). The complete temporal overlap of CS and US (ISI = 0) does not seem to exploit the full potential of learning the association. Instead, partial overlap yields stronger associations due to the extended window of opportunity for synaptic change triggered by the odor's eligibility trace. In our model, the eligibility trace $e(t)$ represents a molecular process that maintains the odor signal locally in the KC>MBON synapses (Equation 2). Appetitive and aversive trace conditioning experiments have been conducted with larvae[34,80,81] and adult flies and other insects.[89,97–99] All experiments during which the CS is presented before the US demonstrate that longer ISIs abolish learning of the CS-US association when no KC odor representation persists during the reward period. In the cases of shorter intervals, the experimental data are not entirely conclusive. The odor preference was either higher for partial or no overlap, compared with complete overlap,[34,99] or highest for complete overlap.[58,81,89]

We also looked at the extent of reward generalization to novel odors. Experiments with larvae have shown that associations between an odor and reward generalize, to a varying extent, to other odors.[79,100] Previous modeling experiments in adult insects have also shown that reward generalization depends on odor similarity.[56,101–103] In our larval model, we also demonstrate both generalization to other odors and a loss in strength, compared to the training odor (Figure 3A). We also show that the extent of the generalization depends on the similarity of the training and test odor, as measured by the overlap of the input patterns (Figure 1D). The larval pathway with its relatively small coding space[62,64,65] might be especially prone to such poor discriminative abilities.

Model predictions for behavioral experiments

We targeted two hypotheses. Firstly, symmetrical inhibitory and excitatory feedback from MBONs to DANs should yield a circuit capable of producing saturating learning curves due to PE[16] driving the learning process, which has also been suggested by previous models.[28,55–58] Secondly, saturating learning curves, driven by PE, should translate into real animal and simulated behavior when comparing different training durations and intensities of reinforcement. We were able to test these hypotheses on the level of MB readout (behavioral bias, Equation 5; Figures 2 and 4) and through the comparison of animal and simulated behavior of artificial larvae (Figure 5). While the simulation results fit nicely with the real larval behavior in an experiment with a varied training duration[59] (Figure 5), ultimately, the role of MBON>DAN feedback needs to be tested in behavioral animal experiments, directly manipulating this feedback. We would expect that the learning curves of individual animals should saturate over time when MBON>DAN feedback is intact. If this feedback was blocked, learning curves should not saturate. Additionally, increasing or decreasing the intensity of the reward should lead to saturation on a higher or lower level, respectively.

Models of PE coding in the insect MB

Models in which a PE signal drives synaptic plasticity lend themselves to studying the evolution and saturation of learning over time (either across repeated learning trials or in continuous time), depending on the learning history. Here, we suggest a minimal basic circuit motif establishing a mechanism for PE coding in the *Drosophila* larva MB. It is based on a two-factor learning rule that requires the coincidence of presynaptic sensory KC activity and DAN activity to reduce the strength of specific KC>MBON synapses and a symmetrical MBON>DAN feedback circuit to convey excitatory and inhibitory feedback components to DANs that compute a PE (Figure 1A). Our model is in line with several recent modeling approaches that have targeted the idea of PE coding in DANs in the adult *Drosophila*[55–58] as well as in the larval MB[28] implemented as some form of MBON>DAN feedback.

Some phenomena predicted by the PE coding theory have been demonstrated in these models. The first key prediction is the saturation of the learning curve across time as the PE decreases,[15,16] which has been shown in trial-based rate models.[55–58] The second prediction is the loss of an acquired association of a stimulus with reinforcement, following repeated presentations of the stimulus alone, which has been realized in two conceptually distinct ways. A memory decay, implemented as a process of changing the KC>MBON weights in the opposite direction of

the learning process, is implemented in our model (Equation 3) and has been used by others in similar ways.[58,103,104] Alternative models have suggested a circuit for extinction learning that can form a parallel extinction memory of opposite valence in the adult[56–58] and the larva.[28] This typically involves an additional cross-compartment inhibitory MBON>DAN feedback and allows for the replication[56] of extinction experiments in the adult fly.[45,72] An additional feature observed in associative learning is trace conditioning, when the reinforcing stimulus is delayed with respect to the CS; it has been realized by an action potential-triggered eligibility trace in our model (Equation 2) and previously through a decaying rate function.[28,57,105]

To our knowledge, the present model is the first implementation of PE coding in the MB in a fully spiking neural network model. The synaptic learning rule (Equation 2) detects a coincidence between a DAN action potential and positive synaptic eligibility that is triggered by KC action potentials. Postsynaptic MBON action potentials trigger a homeostatic restoration of initial weights and thus implement memory loss without reinforcement. The fact that the spiking model is continuous in time allowed us to train and test our model in realistic time-continuous virtual learning experiments and to assess the dynamic change in the model's odor preference. The combination with a time-continuous behavioral simulation[60] during memory retention allowed for straightforward comparison with larval experiments (Figure 5).

Comparison with other insect learning models

Several other models of MB learning, based on KC>MBON plasticity without a PE coding mechanism, exist for the fruit fly,[101,104,106] the honeybee,[102,107] and more general insect MB circuits.[108–110] All of these models suggest plasticity to be mediated by the activity of modulatory neurons, coinciding with either KC[102,103] or coordinated KC and MBON activity.[101,104,108] These models can perform associative learning of a stimulus when paired with reinforcement,[101–104,108] as well as more complicated forms of learning such as second-order conditioning[103,104,111] and matching to sample[108] or reinforcement generalization tasks, the extent of which depends on the stimulus similarity.[101,102] Additionally, some models[101,102,105] were successfully tested in patterning tasks,[112] where combinations of stimuli are reinforced, while their components are not (positive patterning) or vice versa (negative patterning).

Gkanias et al.[103] explored an alternative mechanism to PE for the saturation of learning across trials. Their comprehensive rate model of learning introduced a dopamine-dependent plasticity rule that, similar to Equation 4, consists of a learning term and a recovery term that accounts for memory loss. Their learning rule is a special case of Equation 2 with unity learning rate. Their synaptic recovery term, however, is fundamentally different from our synaptic homeostasis term as it depends on the activity of the reinforcing DAN. Thus, the presence of reward or punishment is required and sufficient to induce loss of memory in contrast to our model, where olfactory input alone is required and sufficient to induce memory loss.

Individuality is comparable for model and animal behavior

Since individual differences in behaviorally expressed learning performance have been observed in insect experiments,[113–118] we included two sources of individuality in our MB model to account for inter-individual variability between model instances. Firstly, the connectivity from PNs to KCs is randomly drawn for each model instance. Secondly, the sensory input to the ORNs and the reward input to DANs are simulated by stochastic point processes. This is sufficient to introduce variability between the learning curves of individuals in Figure 2D with a considerable range of asymptotic values for the behavioral bias. In the behavioral simulations, each simulated larva was placed in the center of the virtual dish with a random head orientation, introducing another source of inter-individual variation. This matches the animal experiments where larvae are placed in the center of the dish using a small brush, resulting in a random head orientation. In addition, the behavioral model implements a stochastic duration for the successive activity and rest bouts.[60] We compare experimental and simulated behavioral results across repeated group experiments by computing the preference index (Equation 17) for each group of 30 larvae at the end of the preference test. The distribution across performance indices (Equation 18) computed from independent groups is slightly lower across simulated groups than across the experimental groups (Figure 5). Additionally, parameters of the larval body could be considered to enhance inter-individual differences, such as the larval size, which affects speed.[60,119,120]

Limitations of the study

Insect experiments have provided mixed evidence for other phenomena that can be predicted from the RW model and PE theory, such as blocking[121–125] and conditioned inhibition.[126–129] We excluded such more complex forms of learning from our experiments. Exploring them in models could yield valuable insights into the *Drosophila* circuit, as well as aid in our general understanding of PE coding. Furthermore, the simplification of the output circuitry to two MBONs prevents the exploration of parallel associations regarding the same stimulus,[45] effectively limiting conclusions about extinction and forgetting. Another model previously demonstrated that computation across compartments could be a fruitful way of approaching PE coding in the MB.[57]

An indirect prediction by the RW model[15] fits the experimental observation of second-order conditioning in adult *Drosophila*,[130–133] where a second CS2 is paired with the CS, after this CS has acquired an association with the US. Through the CS2-CS pairing without the US, the CS2 acquires predictive power of the US. Different mechanisms have been proposed to be involved in causing this effect.[27,103,132,134,111] Currently, second-order conditioning and other more complex forms of learning cannot be accounted for by our model.

CS and US pre-exposure effects[135–138] that might be explained by changes in either attention to the CS or habituation to the CS or the US, caused by prolonged exposure before training, rather than changes in associative strength (for a review see Lubow et al.[139]), are likely intertwined with associative learning in producing animal behavior. They are not addressed by the RW model and are not part of our implementation thereof.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Network model
  - Sparse odor representation
  - Sensory input
  - Experimental protocols
  - Realistic modeling of larval locomotion
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Behavioral bias
  - Spatially-defined measure of learning

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108640.

ACKNOWLEDGMENTS

This project is funded in parts by the German Research Foundation (DFG) within the Research Unit "Structure, Plasticity and Behavioral Function of the Drosophila mushroom body" (DFG-FOR-2705, grant no. 403329959, https://www.unigoettingen.de/en/601524.html to B.G. and M.P.N.) and by the Federal Ministry of Education and Research (BMBF, grant no. 01GQ2103A, "DrosoExpect" to B.G. and M.P.N.). A.-M.J. received additional support from the Research Training Group "Neural Circuit Analysis" (DFG-RTG 1960, grant no. 233886668). P.S. received funding from the Research Training Group "Neural Circuit Analysis" (DFG-RTG 1960, grant no. 233886668). We thank three anonymous reviewers for their valuable comments that helped us improve this manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization, A.-M.J. and M.P.N.; computational experiments & data analyses, A.-M.J. and P.S.; animal data contribution, M.S. and B.G.; writing – original draft, A.-M.J., P.S., and M.P.N.; writing – review & editing, A.-M.J, P.S., M.S., B.G., and M.P.N.; funding acquisition, M.P.N. and B.G.; supervision, M.P.N.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: March 1, 2023
Revised: November 10, 2023
Accepted: December 1, 2023
Published: December 26, 2023

REFERENCES

1. Raymond, J.L., and Lisberger, S.G. (1998). Neural learning rules for the vestibulo-ocular reflex. J. Neurosci. *18*, 9112–9129.
2. Ohmae, S., and Medina, J.F. (2015). Climbing fibers encode a temporal-difference prediction error during cerebellar learning in mice. Nat. Neurosci. *18*, 1798–1803.
3. Sarkisov, D.V., and Wang, S.S.-H. (2008). Order-dependent coincidence detection in cerebellar Purkinje neurons at the inositol trisphosphate receptor. J. Neurosci. *28*, 133–142.
4. Fanselow, M.S., and Poulos, A.M. (2005). The Neuroscience of Mammalian Associative Learning. Annu. Rev. Psychol. *56*, 207–234.
5. Heisenberg, M. (1998). What do the mushroom bodies do for the insect brain? An introduction. Learn. Mem. *5*, 1–10.
6. Stopfer, M. (2014). Central processing in the mushroom bodies. Curr. Opin. Insect Sci. *6*, 99–103.
7. Davis, R.L. (1993). Mushroom bodies and drosophila learning. Neuron *11*, 1–14.
8. Dubnau, J. (2012). Ode to the Mushroom Bodies. Science *335*, 664–665.

9. Heisenberg, M. (2003). Mushroom body memoir: from maps to models. Nat. Rev. Neurosci. *4*, 266–275.
10. Menzel, R. (2012). The honeybee as a model for understanding the basis of cognition. Nat. Rev. Neurosci. *13*, 758–768.
11. Hancock, C.E., Rostami, V., Rachad, E.Y., Deimel, S.H., Nawrot, M.P., and Fiala, A. (2022). Visualization of learning-induced synaptic plasticity in output neurons of the Drosophila mushroom body -lobe. Sci. Rep. *12*, 10421.
12. Aso, Y., Hattori, D., Yu, Y., Johnston, R.M., Iyer, N.A., Ngo, T.T.B., Dionne, H., Abbott, L.F., Axel, R., Tanimoto, H., and Rubin, G.M. (2014b). The neuronal architecture of the mushroom body provides a logic for associative learning. Elife *3*, e04577.
13. Hige, T., Aso, Y., Modi, M.N., Rubin, G.M., and Turner, G.C. (2015). Heterosynaptic plasticity underlies aversive olfactory learning in Drosophila. Neuron *88*, 985–998.
14. Owald, D., and Waddell, S. (2015). Olfactory learning skews mushroom body output pathways to steer behavioral choice in Drosophila. Curr. Opin. Neurobiol. *35*, 178–184.
15. Rescorla, R.A., and Wagner, A.R. (1972). A theory of Pavlovian conditioning : Variations in the effectiveness of reinforcement and non- reinforcement. In Classical Conditioning 2: Current Theory and Research, W.F. Prokasy and A.H. Black, eds. (Appelton-century-Crofts), pp. 64–99. Chap. 3.
16. Kamin, L. (1969). In Predictability, surprise, attention and conditioning. Punsihment and aversive behavior, B.A. Campbell, ed. (Appleton-Century-Crofts), pp. 279–298.
17. Glimcher, P.W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. Proc. Natl. Acad. Sci. USA *108*, 15647–15654.
18. Balsam, P.D., and Gallistel, C.R. (2009). Temporal maps and informativeness in associative learning. Trends Neurosci. *32*, 73–78.
19. Kaplan, P.S. (1984). Importance of relative temporal parameters in trace autoshaping: From excitation to inhibition. J. Exp. Psychol. Anim. Behav. Process. *10*, 113–126.
20. Schultz, W. (1986). Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. J. Neurophysiol. *56*, 1439–1461.
21. Morris, G., Arkadir, D., Nevet, A., Vaadia, E., and Bergman, H. (2004). Coincident but Distinct Messages of Midbrain Dopamine and Striatal Tonically Active Neurons. Neuron *43*, 133–143.
22. Takikawa, Y., Kawagoe, R., and Hikosaka, O. (2004). A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward mapping. J. Neurophysiol. *92*, 2520–2529.
23. Satoh, T., Nakai, S., Sato, T., and Kimura, M. (2003). Correlated Coding of Motivation and Outcome of Decision by Dopamine Neurons. J. Neurosci. *23*, 9913–9923.
24. Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature *482*, 85–88.
25. Deutch, A.Y., Tam, S.-Y., and Roth, R.H. (1985). Footshock and conditioned stress increase 3, 4-dihydroxyphenylacetic acid (DOPAC) in the ventral tegmental area but not substantia nigra. Brain Res. *333*, 143–146.
26. Waddell, S. (2013). Reinforcement signalling in Drosophila; dopamine does it all after all. Curr. Opin. Neurobiol. *23*, 324–329.
27. Riemensperger, T., Völler, T., Stock, P., Buchner, E., and Fiala, A. (2005). Punishment prediction by dopaminergic neurons in Drosophila. Curr. Biol. *15*, 1953–1960.
28. Eschbach, C., Fushiki, A., Winding, M., Schneider-Mizell, C.M., Shao, M., Arruda, R., Eichler, K., Valdes-Aleman, J., Ohyama, T., Thum, A.S., et al. (2020). Recurrent architecture for adaptive regulation of learning in the insect brain. Nat. Neurosci. *23*, 544–555.
29. Lin, S., Owald, D., Chandra, V., Talbot, C., Huetteroth, W., and Waddell, S. (2014). Neural correlates of water reward in thirsty Drosophila. Nat. Neurosci. *17*, 1536–1542.
30. Corbett, D., and Wise, R.A. (1980). Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: A moveable electrode mapping study. Brain Res. *185*, 1–15.
31. Stauffer, W.R., Lak, A., Yang, A., Borel, M., Paulsen, O., Boyden, E.S., and Schultz, W. (2016). Dopamine Neuron-Specific Optogenetic Stimulation in Rhesus Macaques. Cell *166*, 1564–1571.e6.
32. Wise, R.A., and Rompre, P.-P. (1989). Brain dopamine and reward. Annu. Rev. Psychol. *40*, 191–225.
33. Witten, I.B., Steinberg, E.E., Lee, S.Y., Davidson, T.J., Zalocusky, K.A., Brodsky, M., Yizhar, O., Cho, S.L., Gong, S., Ramakrishnan, C., et al. (2011). Recombinase-driver rat lines: Tools, techniques, and optogenetic application to dopamine-mediated reinforcement. Neuron *72*, 721–733.
34. Saumweber, T., Rohwedder, A., Schleyer, M., Eichler, K., Chen, Y.-C., Aso, Y., Cardona, A., Eschbach, C., Kobler, O., Voigt, A., et al. (2018). Functional architecture of reward learning in mushroom body extrinsic neurons of larval Drosophila. Nat. Commun. *9*, 1104.
35. Schleyer, M., Weiglein, A., Thoener, J., Strauch, M., Hartenstein, V., Kantar Weigelt, M., Schuller, S., Saumweber, T., Eichler, K., Rohwedder, A., et al. (2020). Identification of dopaminergic neurons that can both establish associative memory and acutely terminate its behavioral expression. J. Neurosci. *40*, 5990–6006.
36. Schroll, C., Riemensperger, T., Bucher, D., Ehmer, J., Völler, T., Erbguth, K., Gerber, B., Hendel, T., Nagel, G., Buchner, E., and Fiala, A. (2006). Light-induced activation of distinct modulatory neurons triggers appetitive or aversive learning in Drosophila larvae. Curr. Biol. *16*, 1741–1747.
37. Liu, C., Plaaais, P.Y., Yamagata, N., Pfeiffer, B.D., Aso, Y., Friedrich, A.B., Siwanowicz, I., Rubin, G.M., Preat, T., and Tanimoto, H. (2012). A subset of dopamine neurons signals reward for odour memory in Drosophila. Nature *488*, 512–516.
38. Claridge-Chang, A., Roorda, R.D., Vrontou, E., Sjulson, L., Li, H., Hirsh, J., and Miesenböck, G. (2009). Writing Memories with Light-Addressable Reinforcement Circuitry. Cell *139*, 405–415.
39. König, C., Khalili, A., Ganesan, M., Nishu, A.P., Garza, A.P., Niewalda, T., Gerber, B., Aso, Y., and Yarali, A. (2018). Reinforcement signaling of punishment versus relief in fruit flies. Learn. Mem. *25*, 247–257.
40. Aso, Y., and Rubin, G.M. (2016). Dopaminergic neurons write and update memories with cell-typespecific rules. Elife *5*, e16135.
41. Schultz, W., Apicella, P., and Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. J. Neurosci. *13*, 900–913.
42. Schultz, W., Dayan, P., and Montague, P.R. (1997). Getting formal with dopamine and reward. Science *275*, 1593–1599.
43. Zaghloul, K.A., Blanco, J.A., Weidemann, C.T., McGill, K., Jaggi, J.L., Baltuch, G.H., and Kahana, M.J. (2009). Human substantia nigra neurons encode unexpected financial rewards. Science *323*, 1496–1499.
44. Pan, W.-X., Schmidt, R., Wickens, J.R., and Hyland, B.I. (2005). Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. J. Neurosci. *25*, 6235–6242.
45. Felsenberg, J., Jacob, P.F., Walker, T., Barnstedt, O., Edmondson-Stait, A.J., Pleijzier, M.W., Otto, N., Schlegel, P., Sharifi, N., Perisse, E., et al. (2018). Integration of parallel opposing memories underlies memory extinction. Cell *175*, 709–722.e15.
46. Mizunami, M., Terao, K., and Alvarez, B. (2018). Application of a prediction error theory to Pavlovian conditioning in an insect. Front. Psychol. *9*, 1272.
47. Villar, M.E., Pavão-Delgado, M., Amigo, M., Jacob, P.F., Merabet, N., Pinot, A., Perry, S.A., Waddell, S., and Perisse, E. (2022). Differential coding of absolute and relative aversive value in the Drosophila brain. Curr. Biol. *32*, 4576–4592.e5.
48. Rajagopalan, A.E., Darshan, R., Hibbard, K.L., Fitzgerald, J.E., and Turner, G.C. (2023). Reward expectations direct learning and drive operant matching in Drosophila. Proc. Natl. Acad. Sci. USA *120*, e2221415120.
49. Kadow, I.C.G., and Owald, D. (2022). Decision making: Dopaminergic neurons for better or worse. Curr. Biol. *32*, R1237–R1240.
50. Ichinose, T., Aso, Y., Yamagata, N., Abe, A., Rubin, G.M., and Tanimoto, H. (2015). Reward signal in a recurrent circuit drives appetitive long-term memory formation. Elife *4*, e10719.
51. Scheffer, L.K., Xu, C.S., Januszewski, M., Lu, Z., Takemura, S.-Y., Hayworth, K.J., Huang, G.B., Shinomiya, K., Maitlin-Shepard, J., Berg, S., et al. (2020). A connectome and analysis of the adult Drosophila central brain. Elife *9*, e57443.
52. Otto, N., Pleijzier, M.W., Morgan, I.C., Edmondson-Stait, A.J., Heinz, K.J., Stark, I., Dempsey, G., Ito, M., Kapoor, I., Hsu, J., et al. (2020). Input connectivity reveals additional heterogeneity of dopaminergic reinforcement in Drosophila. Curr. Biol. *30*, 3200–3211.e8.
53. Winding, M., Pedigo, B.D., Barnes, C.L., Patsolic, H.G., Park, Y., Kazimiers, T., Fushiki, A., Andrade, I.V., Khandelwal, A., Valdes-Aleman, J., et al. (2023). The connectome of an insect brain. Science *379*, eadd9330.
54. Li, F., Lindsey, J.W., Marin, E.C., Otto, N., Dreher, M., Dempsey, G., Stark, I., Bates, A.S., Pleijzier, M.W., Schlegel, P., et al. (2020). The connectome of the adult Drosophila mushroom body provides insights into function. Elife *9*, e62576.
55. Bennett, J.E.M., Philippides, A., and Nowotny, T. (2021). Learning with

reinforcement prediction errors in a model of the Drosophila mushroom body. Nat. Commun. *12*, 2569.
56. Springer, M., and Nawrot, M.P. (2021). A mechanistic model for reward prediction and extinction learning in the fruit fly. Eneuro *8*, ENEURO.0549-20.2021.
57. Jiang, L., and Litwin-Kumar, A. (2021). Models of heterogeneous dopamine signaling in an insect learning and memory center. PLoS Comput. Biol. *17*, e1009205.
58. Zhao, C., Widmer, Y.F., Diegelmann, S., Petrovici, M.A., Sprecher, S.G., and Senn, W. (2021). Predictive olfactory learning in Drosophila. Sci. Rep. *11*, 6795.
59. Weiglein, A., Gerstner, F., Mancini, N., Schleyer, M., and Gerber, B. (2019). One-trial learning in larval Drosophila. Learn. Mem. *26*, 109–120.
60. Sakagiannis, P., Jürgensen, A.M., and Nawrot, M.P. (2021). A realistic locomotory model of Drosophila larva for behavioral simulations. Preprint at bioRxiv. https://doi.org/10.1101/2021.07.07.451470.
61. Jürgensen, A.M., Khalili, A., Chicca, E., Indiveri, G., and Nawrot, M.P. (2021). A neuromorphic model of olfactory processing and sparse coding in the Drosophila larva brain. Neuromorph. Comput. Eng. *1*, 024008.
62. Berck, M.E., Khandelwal, A., Claus, L., Hernandez-Nunez, L., Si, G., Tabone, C.J., Li, F., Truman, J.W., Fetter, R.D., Louis, M., et al. (2016). The wiring diagram of a glomerular olfactory system. Elife *5*, e14859.
63. Eichler, K., Li, F., Litwin-Kumar, A., Park, Y., Andrade, I., Schneider-Mizell, C.M., Saumweber, T., Huser, A., Eschbach, C., Gerber, B., et al. (2017). The complete connectome of a learning and memory centre in an insect brain. Nature *548*, 175–182.
64. Couto, A., Alenius, M., and Dickson, B.J. (2005). Molecular, anatomical, and functional organization of the Drosophila olfactory system. Curr. Biol. *15*, 1535–1547.
65. Vosshall, L.B., and Stocker, R.F. (2007). Molecular architecture of smell and taste in Drosophila. Annu. Rev. Neurosci. *30*, 505–533.
66. Eschbach, C., Fushiki, A., Winding, M., Afonso, B., Andrade, I.V., Cocanougher, B.T., Eichler, K., Gepner, R., Si, G., Valdes-Aleman, J., et al. (2021). Circuits for integrating learned and innate valences in the insect brain. Elife *10*, e62567.
67. Michels, B., Chen, Y.-C., Saumweber, T., Mishra, D., Tanimoto, H., Schmid, B., Engmann, O., and Gerber, B. (2011). Cellular site and molecular mode of synapsin action in associative learning. Learn. Mem. *18*, 332–344.
68. Aso, Y., Sitaraman, D., Ichinose, T., Kaun, K.R., Vogt, K., Belliart-Guérin, G., Plaçais, P.Y., Robie, A.A., Yamagata, N., Schnaitmann, C., et al. (2014a). Mushroom body output neurons encode valence and guide memory-based action selection in Drosophila. Elife *3*, e04580.
69. Séjourné, J., Plaçais, P.Y., Aso, Y., Siwanowicz, I., Trannoy, S., Thoma, V., Tedjakumala, S.R., Rubin, G.M., Tchénio, P., Ito, K., et al. (2011). Mushroom body efferent neurons responsible for aversive olfactory memory retrieval in Drosophila. Nat. Neurosci. *14*, 903–910.
70. Schwaerzel, M., Heisenberg, M., and Zars, T. (2002). Extinction antagonizes olfactory memory at the subcellular level. Neuron *35*, 951–960.
71. Lesar, A., Tahir, J., Wolk, J., and Gershow, M. (2021). Switch-like and persistent memory formation in individual Drosophila larvae. Elife *10*, e70317.
72. Felsenberg, J., Barnstedt, O., Cognigni, P., Lin, S., and Waddell, S. (2017). Re-evaluation of learned information in Drosophila. Nature *544*, 240–244.
73. Wang, L., Yang, Q., Lu, B., Wang, L., Zhong, Y., and Li, Q. (2019). A behavioral paradigm to study the persistence of reward memory extinction in Drosophila. J. Genet. Genom. *46*, 599–601.
74. Turrigiano, G.G. (2017). The dialectic of Hebb and homeostasis. Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*, 20160258.
75. Pozo, K., and Goda, Y. (2010). Unraveling mechanisms of homeostatic synaptic plasticity. Neuron *66*, 337–351.
76. Vitureira, N., Letellier, M., and Goda, Y. (2012). Homeostatic synaptic plasticity: from single synapses to neural circuits. Curr. Opin. Neurobiol. *22*, 516–521.
77. Rabinowitch, I., and Segev, I. (2008). Two opposing plasticity mechanisms pulling a single synapse. Trends Neurosci. *31*, 377–383.
78. Myers, K.M., and Davis, M. (2002). Behavioral and neural analysis of extinction. Neuron *36*, 567–584.
79. Chen, Y.-C., Mishra, D., Schmitt, L., Schmuker, M., and Gerber, B. (2011). A behavioral odor similarity "space" in larval Drosophila. Chem. Senses *36*, 237–249.
80. Thoener, J., Weiglein, A., Gerber, B., and Schleyer, M. (2022). Optogenetically induced reward and 'frustration'memory in larval Drosophila melanogaster. J. Exp. Biol. *225*, jeb244565.
81. Weiglein, A., Thoener, J., Feldbruegge, I., Warzog, L., Mancini, N., Schleyer, M., and Gerber, B. (2021). Aversive teaching signals from individual dopamine neurons in larval Drosophila show qualitative differences in their temporal "fingerprint". J. Comp. Neurol. *529*, 1553–1570.
82. Saumweber, T., Husse, J., and Gerber, B. (2011). Innate attractiveness and associative learnability of odors can be dissociated in larval Drosophila. Chem. Senses *36*, 223–235.
83. Schleyer, M., Miura, D., Tanimura, T., and Gerber, B. (2015). Learning the specific quality of taste reinforcement in larval Drosophila. Elife *4*, e04711.
84. Michels, B., Diegelmann, S., Tanimoto, H., Schwenkert, I., Buchner, E., and Gerber, B. (2005). A role for Synapsin in associative learning: the Drosophila larva as a study case. Learn. Mem. *12*, 224–231.
85. Bouton, M.E. (2004). Context and behavioral processes in extinction. Learn. Mem. *11*, 485–494.
86. Hirano, Y., Ihara, K., Masuda, T., Yamamoto, T., Iwata, I., Takahashi, A., Awata, H., Nakamura, N., Takakura, M., Suzuki, Y., et al. (2016). Shifting transcriptional machinery is required for long-term memory maintenance and modification in Drosophila mushroom bodies. Nat. Commun. *7*, 13471.
87. Mancini, N., Hranova, S., Weber, J., Weiglein, A., Schleyer, M., Weber, D., Thum, A.S., and Gerber, B. (2019). Reversal learning in Drosophila larvae. Learn. Mem. *26*, 424–435.
88. Eisenhardt, D., and Menzel, R. (2007). Extinction learning, reconsolidation and the internal reinforcement hypothesis. Neurobiol. Learn. Mem. *87*, 167–173.
89. Tully, T., and Quinn, W.G. (1985). Classical conditioning and retention in normal and mutant Drosophila melanogaster. J. Comp. Physiol. *157*, 263–277.
90. Neuser, K., Husse, J., Stock, P., and Gerber, B. (2005). Appetitive olfactory learning in Drosophila larvae: Effects of repetition, reward strength, age, gender, assay type and memory span. Anim. Behav. *69*, 891–898.
91. Mathew, D., Martelli, C., Kelley-Swift, E., Brusalis, C., Gershow, M., Samuel, A.D.T., Emonet, T., and Carlson, J.R. (2013). Functional diversity among sensory receptors in a Drosophila olfactory circuit. Proc. Natl. Acad. Sci. USA *110*, E2134–E2143.
92. Kreher, S.A., Mathew, D., Kim, J., and Carlson, J.R. (2008). Translation of sensory input into behavioral output via an olfactory system. Neuron *59*, 110–124.
93. Cobb, M. (1999). What and how do maggots smell? Biol. Rev. *74*, 425–459.
94. Fishilevich, E., and Vosshall, L.B. (2005). Genetic and functional subdivision of the Drosophila antennal lobe. Curr. Biol. *15*, 1548–1553.
95. Müller, J., Nawrot, M., Menzel, R., and Landgraf, T. (2018). A neural network model for familiarity and context learning during honeybee foraging flights. Biol. Cybern. *112*, 113–126.
96. Twick, I., Lee, J.A., and Ramaswami, M. (2014). Olfactory habituation in Drosophila - odor encoding and its plasticity in the antennal lobe. Prog. Brain Res. *208*, 3–38.
97. Szyszka, P., Demmler, C., Oemisch, M., Sommer, L., Biergans, S., Birnbach, B., Silbering, A.F., and Galizia, C.G. (2011). Mind the gap: olfactory trace conditioning in honeybees. J. Neurosci. *31*, 7229–7239.
98. Galili, D.S., Lüdke, A., Galizia, C.G., Szyszka, P., and Tanimoto, H. (2011). Olfactory trace conditioning in Drosophila. J. Neurosci. *31*, 7240–7248.
99. Tanimoto, H., Heisenberg, M., and Gerber, B. (2004). Event timing turns punishment to reward. Nature *430*, 983.
100. Mishra, D., Louis, M., and Gerber, B. (2010). Adaptive adjustment of the generalization discrimination balance in larval Drosophila. J. Neurogenet. *24*, 168–175.
101. Wessnitzer, J., Young, J.M., Armstrong, J.D., and Webb, B. (2012). A model of non-elemental olfactory learning in Drosophila. J. Comput. Neurosci. *32*, 197–212.
102. Peng, F., and Chittka, L. (2017). A simple computational model of the bee mushroom body can explain seemingly complex forms of olfactory learning and memory. Curr. Biol. *27*, 224–230.
103. Gkanias, E., McCurdy, L.Y., Nitabach, M.N., and Webb, B. (2022). An incentive circuit for memory dynamics in the mushroom body of Drosophila melanogaster. Elife *11*, e75611.
104. Faghihi, F., Moustafa, A.A., Heinrich, R., and Wörgötter, F. (2017). A computational model of conditioning inspired by Drosophila olfactory system. Neural Netw. *87*, 96–108.
105. Haenicke, J. (2015). Modeling insect inspired mechanisms of neural and behavioral plasticity (Freie Universität Berlin).

106. Rapp, H., and Nawrot, M.P. (2020). A spiking neural program for sensorimotor control during foraging in flying insects. Proc. Natl. Acad. Sci. USA *117*, 28412–28421.
107. Häusler, C., Nawrot, M.P., and Schmuker, M. (2011). A spiking neuron classifier network with a deep architecture inspired by the olfactory system of the honeybee. In 5th International IEEE/EMBS Conference on Neural Engineering (IEEE), pp. 198–202.
108. Arena, P., Patané, L., Stornanti, V., Termini, P.S., Zäpf, B., and Strauss, R. (2013). Modeling the insect mushroom bodies: Application to a delayed match-to-sample task. Neural Netw. *41*, 202–211.
109. Huerta, R., and Nowotny, T. (2009). Fast and robust learning by reinforcement signals: Explorations in the insect brain. Neural Comput. *21*, 2123–2151.
110. Schmuker, M., Pfeil, T., and Nawrot, M.P. (2014). A neuromorphic network for generic multivariate data classification. Proc. Natl. Acad. Sci. USA *111*, 2081–2086.
111. Jürgensen, A.M., Schmitt, F.J., and Nawrot, M.P. (2023). Minimal circuit motifs for second-order conditioning in the insect mushroom body. Preprint at bioRxiv. https://doi.org/10.1101/2023.09.11.557174.
112. Deisig, N., Lachnit, H., Giurfa, M., and Hellstern, F. (2001). Configural olfactory learning in honeybees: negative and positive patterning discrimination. Learn. Mem. *8*, 70–78.
113. Chittka, L., Dyer, A.G., Bock, F., and Dornhaus, A. (2003). Bees trade off foraging speed for accuracy. Nature *424*, 388.
114. Chittka, L., and Thomson, J.D. (1997). Sensori-motor learning and its relevance for task specialization in bumble bees. Behav. Ecol. Sociobiol. *41*, 385–398.
115. Muller, H., and Chittka, L. (2012). Consistent interindividual differences in discrimination performance by bumblebees in colour, shape and odour learning tasks. Entomol. Gen. *34*, 1–8.
116. Pamir, E., Szyszka, P., Scheiner, R., and Nawrot, M.P. (2014). Rapid learning dynamics in individual honeybees during classical conditioning. Front. Behav. Neurosci. *8*, 313.
117. Pamir, E., Chakroborty, N.K., Stollhoff, N., Gehring, K.B., Antemann, V., Morgenstern, L., Felsenberg, J., Eisenhardt, D., Menzel, R., and Nawrot, M.P. (2011). Average group behavior does not represent individual behavior in classical conditioning of the honeybee. Learn. Mem. *18*, 733–741.
118. Arican, C., Bulk, J., Deisig, N., and Nawrot, M.P. (2019). Cockroaches show individuality in learning and memory during classical and operant conditioning. Front. Physiol. *10*, 1539.
119. Aleman-Meza, B., Jung, S.-K., and Zhong, W. (2015). An automated system for quantitative analysis of Drosophila larval locomotion. BMC Dev. Biol. *15*, 11.
120. Thane, M., Paisios, E., Stöter, T., Krüger, A.R., Gläß, S., Dahse, A.-K., Scholz, N., Gerber, B., Lehmann, D.J., and Schleyer, M. (2023). High-resolution analysis of individual Drosophila melanogaster larvae uncovers individual variability in locomotion and its neurogenetic modulation. Open Biol. *13*, 220308.
121. Terao, K., and Mizunami, M. (2017). Roles of dopamine neurons in mediating the prediction error in aversive learning in insects. Sci. Rep. *7*, 14694.
122. Smith, B.H., and Cobey, S. (1994). The olfactory memory of the honeybee Apis mellifera. II: Blocking between odorants in binary mixtures. J. Exp. Biol. *195*, 91–108.
123. Thorn, R.S., and Smith, B.H. (1997). The olfactory memory of the honeybee Apis mellifera. III: Bilateral sensory input is necessary for induction and expression of olfactory blocking. J. Exp. Biol. *200*, 2045–2055.
124. Hosler, J.S., and Smith, B.H. (2000). Blocking and the detection of odor components in blends. J. Exp. Biol. *203*, 2797–2806.
125. Gerber, B., and Ullrich, J. (1999). No evidence for olfactory blocking in honeybee classical conditioning. J. Exp. Biol. *202*, 1839–1854.
126. Takeda, K. (1961). Classical conditioned response in the honey bee. J. Insect Physiol. *6*, 168–179.
127. Daly, K.C., and Smith, B.H. (2000). Associative olfactory learning in the moth Manduca sexta. J. Exp. Biol. *203*, 2025–2038.
128. Schleyer, M., Fendt, M., Schuller, S., and Gerber, B. (2018). Associative learning of stimuli paired and unpaired with reinforcement: evaluating evidence from maggots, flies, bees, and rats. Front. Psychol. *9*, 1494.
129. Sen, E., El-Keredy, A., Jacob, N., Mancini, N., Asnaz, G., Widmann, A., Gerber, B., and Thoener, J. (2023). Cognitive Limits of Larval Drosophila: Testing for Conditioned Inhibition, Sensory Preconditioning and Second-Order Conditioning. Preprint at bioRxiv. https://doi.org/10.1101/2023.08.21.554112.
130. Tabone, C.J., and deBelle, J.S. (2011). Second-order conditioning in Drosophila. Learn. Mem. *18*, 250–253.
131. König, C., Khalili, A., Niewalda, T., Gao, S., and Gerber, B. (2019). An optogenetic analogue of second-order reinforcement in Drosophila. Biol. Lett. *15*, 20190084.
132. Yamada, D., Bushey, D., Li, F., Hibbard, K.L., Sammons, M., Funke, J., Litwin-Kumar, A., Hige, T., and Aso, Y. (2023). Hierarchical architecture of dopaminergic circuits enables second-order conditioning in Drosophila. Elife *12*, e79042.
133. Hussaini, S.A., Komischke, B., Menzel, R., and Lachnit, H. (2007). Forward and backward secondorder Pavlovian conditioning in honeybees. Learn. Mem. *14*, 678–683.
134. Rachad, E.Y. (2023). Neural circuit plasticity underlying learning and memory in Drosophila melanogaster: from synaptic connections to behavior. Universität zu Göttingen.
135. Barron, A.B., and Corbet, S.A. (1999). Pre-exposure affects the olfactory response of Drosophila melanogaster to menthol. Entomol. Exp. Appl. *90*, 175–181.
136. Fernández, V.M., Giurfa, M., Devaud, J.-M., and Farina, W.M. (2012). Latent inhibition in an insect: the role of aminergic signaling. Learn. Mem. *19*, 593–597.
137. Jacob, P.F., Vargas-Gutierrez, P., Okray, Z., Vietti-Michelina, S., Felsenberg, J., and Waddell, S. (2021). Prior experience conditionally inhibits the expression of new learning in Drosophila. Curr. Biol. *31*, 3490–3503.e3.
138. Chandra, S.B., Hosler, J.S., and Smith, B.H. (2000). Heritable variation for latent inhibition and its correlation with reversal learning in honeybees (Apis mellifera). J. Comp. Psychol. *114*, 86–97.
139. Lubow, R.E., Weiner, I., and Schnur, P. (1981). In Conditioned Attention Theory. Psychology of Learning and Motivation, *15*, G.H. Bower, ed. (Elsevier), pp. 1–49.
140. Si, G., Kanwal, J.K., Hu, Y., Tabone, C.J., Baron, J., Berck, M., Vignoud, G., and Samuel, A.D.T. (2019). Structured Odorant Response Patterns across a Complete Olfactory Receptor Neuron Population. Neuron *101*, 950–962.e7.
141. Schulze, A., Gomez-Marin, A., Rajendran, V.G., Lott, G., Musy, M., Ahammad, P., Deogade, A., Sharpe, J., Riedl, J., Jarriault, D., et al. (2015). Dynamical feature extraction at the sensory periphery guides chemotaxis. Elife *4*, e06694.
142. Gorur-Shandilya, S., Demir, M., Long, J., Clark, D.A., and Emonet, T. (2017). Olfactory receptor neurons use gain control and complementary kinetics to encode intermittent odorant stimuli. Elife *6*, e27670.
143. Nagel, K.I., and Wilson, R.I. (2011). Biophysical mechanisms underlying olfactory receptor neuron dynamics. Nat. Neurosci. *14*, 208–216.
144. Demmer, H., and Kloppenburg, P. (2009). Intrinsic membrane properties and inhibitory synaptic input of Kenyon cells as mechanisms for sparse coding? J. Neurophysiol. *102*, 1538–1550.
145. Kropf, J., and Rössler, W. (2018). In-situ recording of ionic currents in projection neurons and Kenyon cells in the olfactory pathway of the honeybee. PLoS One *13*, e0191425.
146. Farkhooi, F., Froese, A., Muller, E., Menzel, R., and Nawrot, M.P. (2013). Cellular adaptation facilitates sparse and reliable coding in sensory pathways. PLoS Comput. Biol. *9*, e1003251.
147. Betkiewicz, R., Lindner, B., and Nawrot, M.P. (2020). Circuit and cellular mechanisms facilitate the transformation from dense to sparse coding in the insect olfactory system. Eneuro *7*, ENEURO.0305-18.2020.
148. Olshausen, B.A., and Field, D.J. (2004). Sparse coding of sensory inputs. Curr. Opin. Neurobiol. *14*, 481–487.
149. Barlow, H.B. (1959). Sensory mechanisms, the reduction of redundancy, and intelligence. In Symposium proceedings: Mechanisation of thought processes (Her Majesty's Stationery Office), pp. 535–539.
150. Ito, I., Ong, R.C.-Y., Raman, B., and Stopfer, M. (2008). Sparse odor representation and olfactory learning. Nat. Neurosci. *11*, 1177–1184.
151. Herikstad, R., Baker, J., Lachaux, J.P., Gray, C.M., and Yen, S.C. (2011). Natural movies evoke spike trains with low spike time variability in cat primary visual cortex. J. Neurosci. *31*, 15844–15860.
152. Haider, B., Krause, M.R., Duque, A., Yu, Y., Touryan, J., Mazer, J.A., and McCormick, D.A. (2010). Synaptic and Network Mechanisms of Sparse and Reliable Visual Cortical Activity during Nonclassical Receptive Field Stimulation. Neuron *65*, 107–121.
153. Häusler, C., Susemihl, A., and Nawrot, M.P. (2013). Natural image sequences constrain dynamic receptive fields and imply a sparse code. Brain Res. *1536*, 53–67.
154. Martelli, C., and Fiala, A. (2019). Slow presynaptic mechanisms that mediate

adaptation in the olfactory pathway of Drosophila. Elife *8*, e43735.

155. Farkhooi, F., Muller, E., and Nawrot, M.P. (2011). Adaptation reduces variability of the neuronal population code. Phys. Rev. *83*, 050905.
156. Kreher, S.A., Kwon, J.Y., and Carlson, J.R. (2005). The molecular basis of odor coding in the Drosophila larva. Neuron *46*, 445–456.
157. Hoare, D.J., Humble, J., Jin, D., Gilding, N., Petersen, R., Cobb, M., and McCrohan, C. (2011). Modeling peripheral olfactory coding in Drosophila larvae. PLoS One *6*, e22996.
158. Michels, B., Saumweber, T., Biernacki, R., Thum, J., Glasgow, R.D.V., Schleyer, M., Chen, Y.C., Eschbach, C., Stocker, R.F., Toshima, N., et al. (2017). Pavlovian conditioning of larval Drosophila: An illustrated, multilingual, hands-on manual for odor-taste associative learning in maggots. Front. Behav. Neurosci. *11*, 45.
159. Stimberg, M., Brette, R., and Goodman, D.F. (2019). Brian 2, an intuitive and efficient neural simulator. Elife *8*, e47314.
160. Wystrach, A., Lagogiannis, K., and Webb, B. (2016). Continuous lateral oscillations as a core mechanism for taxis in Drosophila larvae. Elife *5*, e15504.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Published animal behavioral data	Weiglein, A., Gerstner, F., Mancini, N., Schleyer, M., & Gerber, B. (2019)[59]. One-trial learning in larval Drosophila. *Learning & Memory*, *26*(4), 109-120.	https://doi.org/10.1101/lm.049106.118, https://learnmem.cshlp.org/content/26/4/109/suppl/DC1
Software and algorithms		
Python 3 packages	Python Software Foundation	https://www.python.org/downloads/
Code generated for this publication	Lab Martin P. Nawrot	https://github.com/nawrotlab/PEcodingDosophilaMB

RESOURCE AVAILABILITY

Lead contact

Further information can be obtained from the lead contact, Martin P. Nawrot (martin.nawrot@uni-koeln.de).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Published animal experimental data[59] is publicly accessible here: https://learnmem.cshlp.org/content/26/4/109/suppl/DC1.
- All code for the model implementation and the simulation experiments is publicly accessible here: https://github.com/nawrotlab/PEcodingDosophilaMB. All code for the simulation of larval locomotion (Larvaworld) is publicly accessible here: https://github.com/nawrotlab/larvaworld.
- Simulated data can be recreated based on the STAR Methods using the publicly accessible code.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

We re-analyzed published behavioral data in larvae[59] (see also key resources table). No additional animal experiments were conducted for the present study. Experiments complied with the applicable law according to the authorities of the state of Saxony-Anhalt, Germany, and the method details can be found in the original publication.[59] In brief, *Drosophila melanogaster* flies of the Canton-S wild-type strain were maintained at 25°C, 60%–70% relative humidity, and a 12/12 h light–dark cycle. For experiments, five-day-old third-instar feeding-stage larvae of either sex were used. There is no evidence that sex affects the assessed behavior. Cohorts of approximately 30 larvae were collected from the food vials, rinsed in water, and used in the respective experiment. For paired training, the larvae were placed at the center of a Petri dish filled with 1% agarose solution supplemented with 2 mol/l fructose as a taste reward (+). The dish was equipped with two custom-made Teflon containers of 5 mm diameter, filled with 10 μL of n-amylacetate diluted 1:20 in paraffin oil. Larvae were allowed to move on this 'odor+' Petri dish for 1, 2, 2.5, 4, or 8 minutes before being transferred to a second Petri dish that lacked fructose and was equipped with two empty odor containers. In this dish, the larvae could move for the same amount of time. After such training with 'odor+' followed by 'no odor' training, they were transferred to the center of a test Petri dish, which may or may not have contained fructose, and where an odor-filled container was presented on one side and an empty container on the opposite side. After the end of this 3min test phase, the number of larvae on the odor side, the side with no odor, and in a 10-mm wide middle zone was counted. The preference index was calculated (Equation 17, quantification and statistical analysis in STAR Methods). Across repetitions of the experiments, in half of the cases, the sequence was as indicated ('odor+' followed by 'no odor'). In the other cases, it was reversed ('no odor' followed by 'odor+'). The procedure for unpaired training was the same, except that the Petri dishes featured either only the odor or only the reward. After such 'odor' followed by 'no odor+' training (again, in half of the cases, the sequence was reversed: 'no odor+' followed by 'odor'), the preference test was carried out as above. From the preference indices after paired and unpaired training, a performance index was calculated according to (Equation 18, quantification and statistical analysis in STAR Methods).

METHOD DETAILS

Network model

All neurons are modeled as leaky integrate-and-fire neurons with conductance-based synapses. They elicit a spike whenever the threshold is crossed (parameters provided in the Table S1). Each neuronal membrane potential v_i is reset to the resting potential V_r whenever a spike occurs. This is followed by an absolute refractory period of 2 ms, during which the neuron does not integrate any inputs. Any neuron from a given population (*v*O, *v*P, *v*L, *v*K, *v*A, *v*M, *v*D) is governed by the respective equation for ORNs, PNs, LNs, KCs, APL, MBONs and DANs (Equations 6, 7, 8, 9, 10, 11, and 12), Figure 1A). Depending on the neuron type, in addition to a leak conductance g_L, the equations consist of excitatory g_e and inhibitory g_i synaptic input. In the case of the DANs, one excitatory $g_e^{M\mp D}(E_E - v_i^D)$ and inhibitory $g_i^{M\pm D}(E_I - v_i^D)$ input represent the two types of MBON feedback for the reward and punishment encoding DAN, respectively. An additional spike-triggered adaptation conductance was implemented for ORNs, KCs, MBONs, and DANs (Equation 13,[61]), in accordance with our current knowledge of the adaptive nature of ORNs in the larva[140] and the adult fly.[141,142,143] Adaptation in KCs has so far only been demonstrated in other insects.[144,145] In the model of these neurons, the adaptation conductance g_{la} is increased with every spike and decays over time with τ_{la} following previous model implementations of ORNs and KCs as leaky integrate-and-fire neurons with spike frequency adaptation (SFA).[47,61,106,146,147] The mechanism of synaptic plasticity is described in the results section (learning and synaptic plasticity at the KC > MBON synapses).

$$C_m\frac{d}{dt}v_i^O = g_L^O\left(E_L^O - v_i^O\right) + g_e^{InputO}\left(E_E - v_i^O\right) - g_{la}\left(E_{la} - v_i^O\right) \quad \text{(Equation 6)}$$

$$C_m\frac{d}{dt}v_i^L = g_L^L\left(E_L^L - v_i^L\right) + g_e^{OL}\left(E_E - v_i^L\right) \quad \text{(Equation 7)}$$

$$C_m\frac{d}{dt}v_i^P = g_L^P\left(E_L^P - v_i^P\right) + g_e^{OP}\left(E_E - v_i^P\right) - g_i^{LP}\left(E_I - v_i^P\right) \quad \text{(Equation 8)}$$

$$C_m\frac{d}{dt}v_i^K = g_L^K\left(E_L^K - v_i^K\right) - g_i^{APLK}\left(E_I - v_i^K\right) + g_e^{PK}\left(E_E - v_i^K\right) - g_{la}\left(E_{la} - v_i^K\right) \quad \text{(Equation 9)}$$

$$C_m\frac{d}{dt}v_i^A = g_L^A\left(E_L^A - v_i^A\right) + g_e^{KAPL}\left(E_E - v_i^A\right) \quad \text{(Equation 10)}$$

$$C_m\frac{d}{dt}v_i^M = g_L^M\left(E_L^M - v_i^M\right) + g_e^{KM}\left(E_E - v_i^M\right) \quad \text{(Equation 11)}$$

$$C_m\frac{d}{dt}v_i^D = g_L^D\left(E_L^D - v_i^D\right) - g_i^{M_\pm D}\left(E_I - v_i^D\right) + g_e^{M_\mp D}\left(E_E - v_i^D\right) + g_e^{InputD}\left(E_E - v_i^D\right) \quad \text{(Equation 12)}$$

$$\frac{d}{dt}g_{la} = -\frac{g_{la}}{\tau_{la}}. \quad \text{(Equation 13)}$$

We based our circuit model on the larval connectome both in terms of connectivity as well as numbers of neurons in each population[28,62,63] and introduced simplifications to support the mechanistic investigation of the MBON>DAN feedback circuit and its role in PE coding and excluded a number of connections that have been demonstrated in the larva. Due to the limited availability of anatomical, functional, and behavioral data, most of our circuit implementation is based on the first instar larva,[28,62,63] while the information on the APL connectivity within the circuit originates from studies on the third instar larva.[34] Behavioral experiments used for comparison with our simulation results were also performed with third instar larvae.[34,59,81] We demonstrate that our model based on the less developed circuit in the first instar larva is sufficient to reproduce animal behavior as observed in the older animals. From the anatomy of the first instar larva we excluded DAN>KC[63] and DAN>MBON synapses[63] that may play an additional role in learning-induced plasticity at KC>MBON synapses,[63] the details of which are not fully known. Instead, we induce plasticity purely via the simulation of a neuromodulatory effect of the DANs onto the KC>MBON synapses.[63] We also neglect recurrent interactions among KC themselves.[63] Many of these interactions affect KC encoding different sensory modalities, which are not included in our exclusively olfactory model. Furthermore, we simplified the connectivity between LNs and PNs[62] and between PNs and KCs to $2 - 6$ PN inputs per KC, which excludes the set of KCs in the larva that receives exclusive input from only one PN.[63] This modification supported model robustness with respect to odor encoding within the small set of 72 KCs. Finally, from the population of $\approx$ 25 larval MBONs, we only modeled two and correspondingly adapted KC>MBON synapses to provide both MBONs with input from all KCs.

Sparse odor representation

We implemented four mechanisms supporting population- and temporal sparseness in the MB odor representation.[61] Population sparseness is defined as the activation of only a small subset of neurons by any given input.[148] In this circuit, population sparseness is enhanced through lateral inhibition (via LNs), inhibitory APL feedback, and the divergent connectivity from PNs to a larger number of KCs.[61] Temporal sparseness

refers to the efficient and economic stimulus encoding in time where a specific stimulus configuration is represented by only a few action potentials,[149–151] which supports encoding of dynamic changes in the sensory environment.[106,152–154] In our model, temporal sparseness is facilitated by the cellular mechanisms of spike frequency adaptation across successive stages of the olfactory pathway, specifically in ORNs and KCs, and by inhibitory feedback via the APL onto the KCs.[61] At the phenomenological level, the larval ORNs have been shown to experience firing rate adaptation in response to abrupt changes in the odor stimulus concentration, and behavioral experiments showed that the ORN adaptation affects the animals' chemotactic behavior when navigating through an odor landscape.[141] In a computational model of the adult fruit fly, temporal sparseness has been shown to support recognition of odor identity in brief odor filament encounters during flight.[106] In addition and on theoretical ground,[146,155] spike frequency adaptation increases neuronal response reliability.

Sensory input

At the periphery of the olfactory pathway of the larval *Drosophila*, any odor activates up to $\approx$ 1/3 of ORNs, depending on odor concentration.[140,156] We implemented receptor input to ORNs by stochastic point processes resulting in postsynaptic currents to mimic the stochastic transduction process at each receptor.[143] Each of the 21 receptor inputs is modeled according to a gamma process (shape parameter $k = 3$). The spontaneous firing rate of larval ORNs has been measured in the range of $0.2 - 7.9$ Hz, depending strongly on odor and receptor type.[141,156,157] ORNs in our model exhibit an average spontaneous firing rate of 8.92 Hz (SD= 0.2 Hz). We constructed realistic olfactory input across the ORN population for amylacetate and 3-octanol by estimating ORN spike frequency from calcium signals measured in the receptor neurons[140] (dilution of 10^{-4} [140]) to ensure that spike rates would not exceed the rates reported by Kreher et al.[156] They showed that using an even stronger odor concentration (dilution 10^{-2}[140]) ORN spiking activity never exceeded a frequency of 200 Hz. Due to the lower concentration used for amylacetate and 3-octanol (Figure 1D)[140] in our experiments and because the authors in[156] estimated firing rates only during the first 0.5s after odor onset when the effect of spike frequency adaptation in ORNs is the weakest (leading to higher spike rates), we decided to allow for a maximum of 150 Hz in odor activated ORNs. After generating the gamma process realizations, we clipped multiple spikes occurring in each time step of the simulation for technical reasons, discarding all but the first spike in each time step. Similar to the odor input, the presence of either reward in the experimental context was implemented as input to the DAN_+/DAN_-. Regular gamma spike trains ($k = 10$) were generated and clipped for the odor input.

To assess the effects of odor similarity on generalization, we additionally created four artificial odors (A, B, C, D, Figure 1D). We quantified the pair-wise distances in the ORN coding space using the cosine distance (Equation 14), where vectors a and b each represent the input spike rate for one odor.

$$D_{cos} = 1 - \frac{\sum_{i=1}^{n} a_i \cdot b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} \cdot \sqrt{\sum_{i=1}^{n} b_i^2}}. \quad \text{(Equation 14)}$$

The cosine distance between odors A and B equals 0.21, 0.77 between odors A and C, and 0.99 between odors A and D. The comparison of amylacetate and 3-octanol yields a distance of 0.16.

Experimental protocols

The simulation experiments reported here belong to one of three categories. The first was performed to provide insight into the model and the effects of specific circuit functions on synaptic plasticity and PE coding. To this end, we used amylacetate as the primary odor input. We varied the intensity of the reward via the frequency of the gamma spike train, provided as input into the DAN_+ (either 500 Hz or 550 Hz, resulting in an average output spike rate of 33.11 Hz / 39.14 Hz), and the learning rate α (0.6 nS or 0.8 nS). Additionally, MBON>DAN feedback was either enabled or disabled (Figure 1A).

Experiments belonging to the second category were designed to replicate larva lab experiments to allow for a direct comparison with our model results. With these comparisons, we aim to validate the model and show to what extent our assumptions about the circuit functions allow us to recreate experimental data (Figure 5). Replicating lab experiments also provides more insights into the circuit mechanisms and offers alternative interpretations of the phenomena observed in data from animal experiments. Our implementations of the lab experiments were set up following the general procedure described in the Maggot Learning Manual.[158] Regardless of the specific protocol used in different experiments, larvae are placed into Petri dishes in groups of 30 animals. They are allowed to move around freely on the substrate that contains reinforcing substances, such as sugar or bitter tastants. During the entire time, they are subjected to specific odorants emitted from two small containers in the dish to create permanent and relatively uniformly distributed odor exposure within the dish. In the analogy of the experimental setting, in our simulated experiments, each model instance is trained individually through the concurrent presentation of olfactory stimulation and reward. One-minute intervals with only baseline ORN stimulation were included between training trials to simulate the time needed in the lab experiments for transferring larvae between Petri dishes. Unless otherwise specified and test phases refer to 3 min, during which only odors are presented. All simulations were implemented in the network simulator Brian2.[159]

Realistic modeling of larval locomotion

Behavior during the testing phase of the olfactory learning experiment was simulated via the freely available python-based simulation platform Larvaworld (https://github.com/nawrotlab/larvaworld,[60]). A group of 30 virtual larvae were placed with random initial orientation around the center of a 100 mm diameter Petri dish and left to behave unconstrainedly for 3 minutes. The previously conditioned odor was placed at

one side of the dish, 10 mm from the arena's boundary. Each larva features a bi-segmental body supervised by a layered control architecture.[60] The basic layer of the control architecture is a locomotory model capable of realistic autonomous exploration behavior. It consists of two coupled oscillators, one representing the crawling apparatus that generates forward velocity oscillations resembling consecutive peristaltic strides.[60] The other oscillator generates alternating left and right lateral bending, manifested as angular velocity oscillations.[160] The crawling and the bending oscillators are coupled via phase-locked suppression of lateral bending to capture the bend dependency on the stride-cycle phase during crawling (weathervaning). Finally, intermittent crawling is achieved by a superimposed intermittency module that generates alternating epochs of crawling and stationary pauses, with more headcasts for orientation during the latter.[60]

Modulation of behavior due to sensory stimulation is introduced at the second reactive layer of the control architecture. An odor signal can transiently alter both the amplitude and frequency of the lateral bending oscillator, which biases free exploration towards approach or avoidance along an olfactory chemical gradient. This modulation of behavior is directly influenced via top-down signaling from the third adaptive layer of the control architecture. In our approach, the spiking MB model populates the adaptive layer and its learning-dependent output, defined as the behavioral bias (i.e., the difference in MBON firing rates, Equation 5), provides the top-down signal.[66] We formalize the gain of behavioral modulation as

$$G = g \cdot BB. \quad \text{(Equation 15)}$$

which is directly proportional to the behavioral bias and the additional proportionality factor $g = 0.5$.

QUANTIFICATION AND STATISTICAL ANALYSIS

Behavioral bias

We computed the difference between MBON spike counts as a measure of odor valence at the MB output. Either for bins of $T = 1$ s to quantify continuous changes during learning (Figure 2)

$$BB = \frac{MBON_{+} - MBON_{-}}{T} \quad \text{(Equation 16)}$$

or across the duration of a $T = 3$ min test phase (Figures 3A, 4, and 5).

Spatially-defined measure of learning

A set of $10 * 30$ trained MB model instances was used to generate 10 groups of 30 simulated larvae. The preference index and the performance index[59] for these simulations are illustrated in Figure 5.

Preference indices (Pref) were computed individually for the paired and the unpaired experiments,[59] based on the number of animals on each side (odor vs. no odor) of the Petri dish at the end of the test phase.

$$Pref = \frac{count_{odor} - count_{no\ odor}}{count_{odor} + count_{no\ odor}} \quad \text{(Equation 17)}$$

The Performance indices (PI) are computed from the preference indices of the paired and unpaired experiments.[59]

$$PI = \frac{Pref_{paired} - Pref_{unpaired}}{2} \quad \text{(Equation 18)}$$