Relative value learning in *Drosophila melanogaster* larvae

Sadniman Rahman[1], Nobuaki K. Tanaka[1,2], Michael Schleyer[1,3]*

1 Graduate School of Life Science, Hokkaido University, 060-0810 Sapporo, Japan

2 Faculty of Science, Hokkaido University, 060-0810 Sapporo, Japan

3 Institute for the Advancement of Higher Education, Hokkaido University, 060-0810 Sapporo, Japan

* Correspondence to: m.schleyer@oia.hokudai.ac.jp

Abstract

The ability to learn from past experiences to inform future decision-making is crucial for humans and animals alike. One question with important implications for adaptive decision-making is whether we learn about the absolute values of cues we encounter (how good or bad?), or about their relative values (how much better or worse than the alternative?). Humans have been shown to use relative value learning, even when it leads to suboptimal decisions. In this study, we ask whether insects employ absolute or relative value learning. Using the larvae of the fruit fly *Drosophila melanogaster*, we designed associative odour-taste learning experiments to distinguish both kinds of learning and find that larvae learn about the relative rather than the absolute values of both rewards and punishments, irrespective of the number and sequence of training trials. This suggests that relative value learning is a facility shared across the animal kingdom from maggots to humans, and can be accomplished even by small insect brains. Given the potential of *D. melanogaster* as a model organism for in-depth neurobiological analyses, our study opens up an opportunity to reveal the detailed mechanisms underlying relative value learning.

Keywords: Associative learning, Classical conditioning, Insect, Value normalization

Introduction

One of the most fundamental functions of our memory is to guide adaptive decision-making [1-3]. From what to eat for dinner to which work project to prioritize, decision-making pervades our daily lives. A wealth of evidence suggests that choice options are evaluated in a relative rather than absolute manner, that is, we rate an option according to how it compares to others [4-6]. In many cases, our previous experiences help us to correctly assess the value of each option and to select the best choice. How exactly value is assigned to a given option during such previous experiences is the subject of intensive research.

A common approach to study this question is associative learning. In the last decade, several studies have shown convincingly that humans apply relative evaluation of rewards and punishments during associative learning. This means that humans learn the value of a reward or punishment relative to other rewards or punishments that were available in a particular training context (how much better or worse?) rather than an absolute value (how good or bad?) that would be independent from the training context. This process is known as relative value learning, adaptive value coding, value normalization or range adaptation [7-13]. This discovery is important for our understanding of economic, value-based decisions because relative value learning can lead to suboptimal choices. For example, an option that had been paired with an objectively low value, but in a context of even lower values, is preferred over another option that had been paired with an objectively higher value, but in a context of even higher values. In other words, humans tend to choose options that have been better than the alternatives in previous experiences, even if this implies choosing the objectively worse of the options at hand [8, 14-15].

The last observation raises the question of the adaptive significance of relative value learning. Proposed advantages include a higher efficiency of saving relative value differences rather than absolute values [16-17], and an increased sensitivity to the most likely outcomes in a given environment [18]. Similar proposals have been made in the context of the optimal foraging and decision-making theory [19-20]: in stable environmental conditions in which the availability of different options is known to the forager and does not change over time, learning absolute values of food sources is of advantage and foragers should accept food sources based on a fixed, absolute threshold. In contrast, in fluctuating environments foragers profit from learning constantly updated values relative to their internal state and previously encountered food sources because it provides flexibility, improves cost efficiency and allows to maximize food intake [21-28].

Relative value learning has been found in mammals and birds, suggesting that the principles of value coding are shared across vertebrates [29-35]. Invertebrates like insects have very differently structured brains that are much smaller and often regarded as too simple to allow for complex tasks. Nevertheless, associative learning has been demonstrated in insects like bees, crickets and flies since decades [36-38], and relative evaluation of food sources has been observed during foraging [28]. Several insect species, including flies, bees and ants, have been demonstrated to be capable of discriminating values of different rewards or punishments and form memories of different strength accordingly [39-44]. However, these studies cannot answer the question whether insects actually learn about, that is, "save" the relative value of a reward or punishment, or the absolute value. A study in grasshoppers showed that insects can learn a reward value relative to their current internal state when being trained at different hunger states [45]. Honeybees have been found to anticipate increasing and decreasing rewards, and thus to learn a reward value relative to previous instances of the same reward [46]. Still, the question remains whether insects compare and learn about the relative values of different rewards and punishments that are present during training, even if this leads to suboptimal decisions, like vertebrates do.

Here, we use the larval stage of the fruit fly *Drosophila melanogaster*. In the past decades, the *D. melanogaster* larva has established itself as useful model organism for neurobehavioural research [47-51]. It combines a small brain with only about 10,000 neurons and a rich toolbox for transgenic manipulations of individual cells with a surprisingly complex behavioural repertoire. This includes appetitive and aversive classical as well as operant conditioning, long-term-memories, decision-making and adaptive outcome expectations [39, 52-61].

We designed a behavioural paradigm to distinguish absolute and relative value learning in classical conditioning. Based on established odour-taste associative experiments [52, 62-64], we systematically ask whether larvae learn about the relative values of rewards as well as punishments and whether this learning depends on the odours presented during the training or the memory retrieval test, the number of training trials and the sequence of training trials. Our findings will help to shed light on how insects encode values and provide an opportunity to investigate the neuronal mechanism of relative value learning in a numerical simple brain.

Materials and Methods

Animals

Throughout the study, we used 4 days old, third instar, feeding-stage Canton-S wild-type strain of *Drosophila melanogaster* larvae. We maintained them on the standard fly food at 25°C with a 12:12 h light–dark cycle.

Three trials fructose learning experiment

The associative learning experiments used in this study are based on a commonly used experimental paradigm [63], with some deviations. We trained larvae with two odours, *n*-amylacetate (AM; CAS: 628-63-7, Tokyo Chemical Industry Co., Ltd, Japan) and 1-octanol (OCT; CAS: 111-87-5, Tokyo Chemical Industry) diluted 1:49 and 1:1 in paraffin oil (CAS: 128-04375, Fujifilm Wako Pure Chemical Corporation, Japan), respectively, associated with different concentrations of fructose (CAS: 127-02765, Fujifilm Wako Pure Chemical). Specifically, we used 0.1 M (+) and 2 M (++) mixed in 2 % agar (CAS: 016-11875, Fujifilm Wako Pure Chemical) as weakly and strongly rewarding substrates, respectively, as well as pure, tasteless agar as neutral situation (∅). Every cohort of larvae was trained with each odour being paired with one of the three different substrates. A total of nine combinations were tested in this experiment, summarized in Table 1.

		Reward paired with AM		
		∅	+	++
Reward paired with OCT	∅	AM∅/OCT∅ (1)	AM+/OCT∅ (4)	AM++/OCT∅ (7)
	+	AM∅/OCT+ (2)	AM+/OCT+ (5)	AM++/OCT+ (8)
	++	AM∅/OCT++ (3)	AM+/OCT++ (6)	AM++/OCT++ (9)

Table 1: Experimental design. AM: *n*-amylacetate, OCT:1-octanol, ∅: neutral substrate, +: weak reward, ++: strong reward. The group numbers in brackets are referred to in the text.

At first, approximately 20 larvae were taken from the food vials, gently washed with tap water and placed at the centre of a training Petri dish (9 cm inner diameter, order #3-1491-01, AS ONE, Japan), filled with one of the three substrates. Two custom-made teflon odour containers filled with the same odour (either AM or OCT) had been placed on opposing sides of the dish just before start of the experiment. The larvae were allowed to move freely over the dish for 2.5 minutes.

Subsequently, they were gently transferred to another training dish, filled with one of the three substrates and with the respective other odour placed, and stayed there for another 2.5 min. We repeated this cycle for 3 times. For example, to train group 6 in Table 1 (AM+/OCT++; Fig. S1a), the larvae were placed in a training dish featuring two odour containers filled with AM and the weak reward (AM+), followed by a second dish featuring OCT and the strong reward (OCT++). In half of the repetitions of the experiment, the sequence was reversed such that the training started with OCT++ followed by AM+. After three training cycles, we immediately tested the animals' preference for AM by placing them at the centre of the test Petri dish which featured on one side an odour container filled with AM and on the opposite side an empty container (EM). All experiments in this study used test dishes filled with 5 mM quinine, unless explicitly mentioned otherwise (for more details, see below). After 3 min, the number of larvae (#) on the AM side, on the EM side, and in a 10 mm wide middle zone was counted to calculate their preference for AM using the following formula. Larvae on the lid were excluded from the analysis.

Preference = (#AM − #EM) / (#AM + #middle + #EM)

Thus, the preference is limited between -1 and +1, with negative scores indicating avoidance and positive scores indicating approach to AM. In some experiments, we calculated a memory score from the preferences of two reciprocally trained groups, following the established procedure [63]. In these cases, always the preference of a group in which OCT had been rewarded and AM not (e.g. AM∅/OCT++) was subtracted from a group with opposite contingency (AM++/OCT∅), divided by two:

Memory Score = $(\text{Preference}_{\text{AM++/OCT}\varnothing} - \text{Preference}_{\text{AM}\varnothing\text{/OCT++}})$ / 2.

Thus, the memory score is limited between -1 and +1, with negative scores indicating aversive, and positive scores indicating appetitive associative memory.

For further statistical analyses, we estimated the subjective value (SV) of each of the three substrates from the preference behaviour of the animals. Towards this end, we used the mean preferences of groups 1, 4 and 7 (see Table 1), as in each of these groups AM was paired with a different substrate while OCT was paired with the neutral substrate. We reasoned that these groups are best suited to estimate the value the animals assign to each substrate, independent of whether they learned the absolute or relative value. Specifically, group 1 (AM∅/OCT∅) did not involve any rewards and thus arguably no associative learning and was therefore used to determine the animals' baseline AM preference after the training procedure but in absence of an associative

memory. Groups 4 (AM+/OCT∅) and 7 (AM++/OCT∅) were then used to estimate the SV of the weak and strong reward, respectively. We calculated the SV of each substrate as follows:

$SV_{neutral} = 0$

$SV_{weak} = \text{Mean Preference}_{AM+/OCT\varnothing} - \text{Mean Preference}_{AM\varnothing/OCT\varnothing}$

$SV_{strong} = \text{Mean Preference}_{AM++/OCT\varnothing} - \text{Mean Preference}_{AM\varnothing/OCT\varnothing}$.

From these SV values, we then calculated the difference in value between the two training substrates:

$\text{Value difference} = SV_{\text{paired with AM}} - SV_{\text{paired with OCT}}$.

Variations of the fructose learning experiment

We performed some variations of the experiment described above, always following the same experimental design with 9 combinations of neutral substrate, weak and strong reward, and always testing for the preference of one odour as described.

- Testing for OCT: in one experiment, the animals were trained exactly as described, but the preference for OCT was measured by presenting OCT instead of AM in the test Petri dish. The preference, SV and value difference were calculated analogously for OCT instead of AM.
- Training with one odour: the experiment was performed while using only one odour during training, replacing the other odour with empty containers (Fig. S1b). Such one-odour training is well established [62]. We used either AM, OCT or undiluted Benzaldehyde (BA; CAS: 025-12206, Fujifilm Wako Pure Chemical). For example, instead of training the animals AM+/OCT++ (group 6 in Table 1), they were trained AM+/EM++. The preference, SV and value difference were calculated analogously for the previously trained odour.
- Training with only one training cycle: the training followed the design in Table 1, but each substrate-odour pairing was only presented a single time before testing for the AM preference in an immediate test. This allowed us to compare two possible sequences: either OCT was presented first (e.g. OCT++ → AM+; Fig. S1c) or AM was presented first (AM+ → OCT++; Fig. S1c'). In an additional experiment, we omitted both odours during the training phase, merely exposing the animals to different sugar concentrations before testing them for their AM preference.

Quinine learning experiment

This experiment was performed in the same way, except using two concentrations of quinine (CAS: 177-00462, Fujifilm Wako Pure Chemical), 0.1 mM (-) and 5 mM (--) as weakly and strongly punishing situations (Table 2). In an analogous experiment, we used 9 different combinations of a neutral substrate (pure agar), reward (2 M fructose) and punishment (5 mM quinine). This included groups of larvae that were trained with one odour being paired with a reward and the other one paired with a punishment (Table 3). Here, we used groups 5-6 to calculate the SV.

Previous studies showed that quinine memory is only expressed in presence of quinine, while the expression of fructose memories is independent of the absence or presence of quinine [53]. Therefore, and to maintain comparability across experiments, all experiments in this study used test Petri dishes filled with 5 mM quinine, unless explicitly mentioned otherwise.

		Punishment paired with AM		
		∅	-	--
Punishment paired with OCT	∅	AM∅/OCT∅ (1)	AM-/OCT∅ (4)	AM--/OCT∅ (7)
	-	AM∅/OCT- (2)	AM-/OCT- (5)	AM--/OCT- (8)
	--	AM∅/OCT-- (3)	AM-/OCT-- (6)	AM--/OCT-- (9)

Table 2: Experimental design of the quinine only experiment. For more details, see Table 1.

		Substrate paired with AM		
		--	∅	++
Substrate paired with OCT	--	AM--/OCT-- (1)	AM∅/OCT-- (4)	AM++/OCT-- (7)
	∅	AM--/OCT∅ (2)	AM∅/OCT∅ (5)	AM++/OCT∅ (8)
	++	AM--/OCT++ (3)	AM∅/OCT++ (6)	AM++/OCT++ (9)

Table 3: Experimental design of the fructose-quinine experiment. For more details, see Table 1.

Taste preference experiment

In the learning experiments, the larvae are often exposed to highly concentrated tastants (fructose or quinine) followed by a much lower concentration of the same taste. It is conceivable that after being exposed to a high concentration, the animals' sensing and/or motivation towards the lower concentration is impaired. To control for this, we placed groups of 20 larvae on a Petri dish with either pure agar or agar containing the highly concentrated tastants (2 M fructose or 5 mM quinine)

for 2.5 minutes. Afterwards, the larvae were transferred to a split Petri dish with pure agar on one side and the low concentration of the same tastants (0.1 M fructose or 0.1 mM quinine) on the other side. After 3 minutes, the number of the animals on either side and in a 10 mm middle zone were counted and the taste preference was calculated:

Fru Preference = (#Fru − #Pure) / (#Fru + #middle + #Pure)

Qui Preference = (#Qui − #Pure) / (#Qui + #middle + #Pure) .

Statistical analyses

Data are displayed as raw data points, overlaid with the mean as middle line and the 95 % confidence interval as error bars. For all statistical tests described below, when multiple comparisons were performed within one analysis, a Bonferroni-Holm correction was applied to keep the experiment-wide error rate below 5 % [65].

To determine whether gradual increases of reward or punishment strength consistently affects odour preferences and memory scores, Spearman Rank Correlations (SRC) were performed, followed by pairwise Mann-Whitney U tests (MWU) if the SRC was significant. SRC were chosen over Pearson correlations as SRC does not assume a linear relationship of the data, only a monotonically increasing or decreasing relationship. Spearman's r_S values are displayed within the figures to allow readers to judge the strength of the correlations (+1 and -1 indicating perfect positive or negative correlations, respectively). In experiments measuring taste preference, we compared groups with MWU tests and determined whether the preferences were significantly different from zero using the Wilcoxon Signed-Rank test (WSR).Graphs, MWU, WSR and SRC tests were made in Statistica 11 (StatSoft, Tulsa, USA).

As the preference scores in our experiments are limited between -1 and 1, their relationship with any experimental parameter is likely to follow a sigmoid curve. Therefore, we fitted a simplified generalized logistic function to the data. The generalized logistic function can be changed by a number of parameters, of which we used three in this study [66]. The function followed the form:

Preference = A_{left} + (A_{right} – A_{left}) / (1 + e^(-B*Value difference))

Here, A_{left} and A_{right} indicate the left and right horizontal asymptotes, respectively, and B the growth rate, that is, the steepness of the function. The parameters of the function for each experiment were determined by nonlinear least square fitting using the "nls" function in R [67],

194 and are reported in Supplementary table S1 along with the raw data. To compare the generalized
195 logistic function with a simpler linear function, we compared the Akaike information criterion
196 (AIC) that balances the likelihood of a fitted model with its complexity. The AIC is regularly used
197 to compare statistical models for the same data set, with the model with the lowest AIC being
198 considered the best fit [68].
199

200 **Results**

201 *Larvae learn about the relative value of rewards*

202 We performed an associative learning experiment that was modified from the most commonly used
203 learning paradigm in *D. melanogaster* [63, 69]. We trained larvae by alternatingly pairing *n*-
204 amylacetate (AM) with a given concentration of fructose reward and 1-octanol (OCT) without
205 reward, and subsequently tested them for their preference of AM alone (Fig. 1a). The presence of
206 only AM during the test, different from the standard procedure, allowed us to assess what the
207 animals had learned specifically about this odour. Two fructose concentrations were chosen based
208 on initial experiments (Fig. S2a-c). We ensured that the larvae can still detect and process the lower
209 of the two fructose concentrations immediately after being exposed to the higher concentration,
210 although their preference was slightly reduced (Fig. S2d).

211 Larvae have a positive innate preference to all odours used in this study that can be modulated
212 by non-associative processes [62, 64]. We therefore included a baseline group with the same
213 handling and odour exposure as all other groups, but without any rewards and therefore arguably
214 without associative learning (Fig. 1a, left-most group). Compared to this baseline, we found that
215 the animals' preference for AM increased with increasing reward concentration, suggesting
216 stronger associative learning (Fig. 1a, see also Fig. S2a-c).

217 Importantly, in this experiment the larvae could learn to associate AM either with an absolute
218 value of the paired reward (how good it was), or with its relative value (how much better it was
219 than the alternative). As the alternative during training was plain, neutral agar, both hypotheses
220 predict the same outcome and the current experiment does not allow to differentiate between them.

221 Therefore, we trained the larvae with AM and OCT being paired with the same fructose
222 concentration (Fig. 1b). If the animals learned about the absolute reward value, the substrate that
223 was paired with OCT should not matter and the animals should show the same level of associative
224 learning as in Fig. 1a. If the animals learned about the relative reward value, however, there should

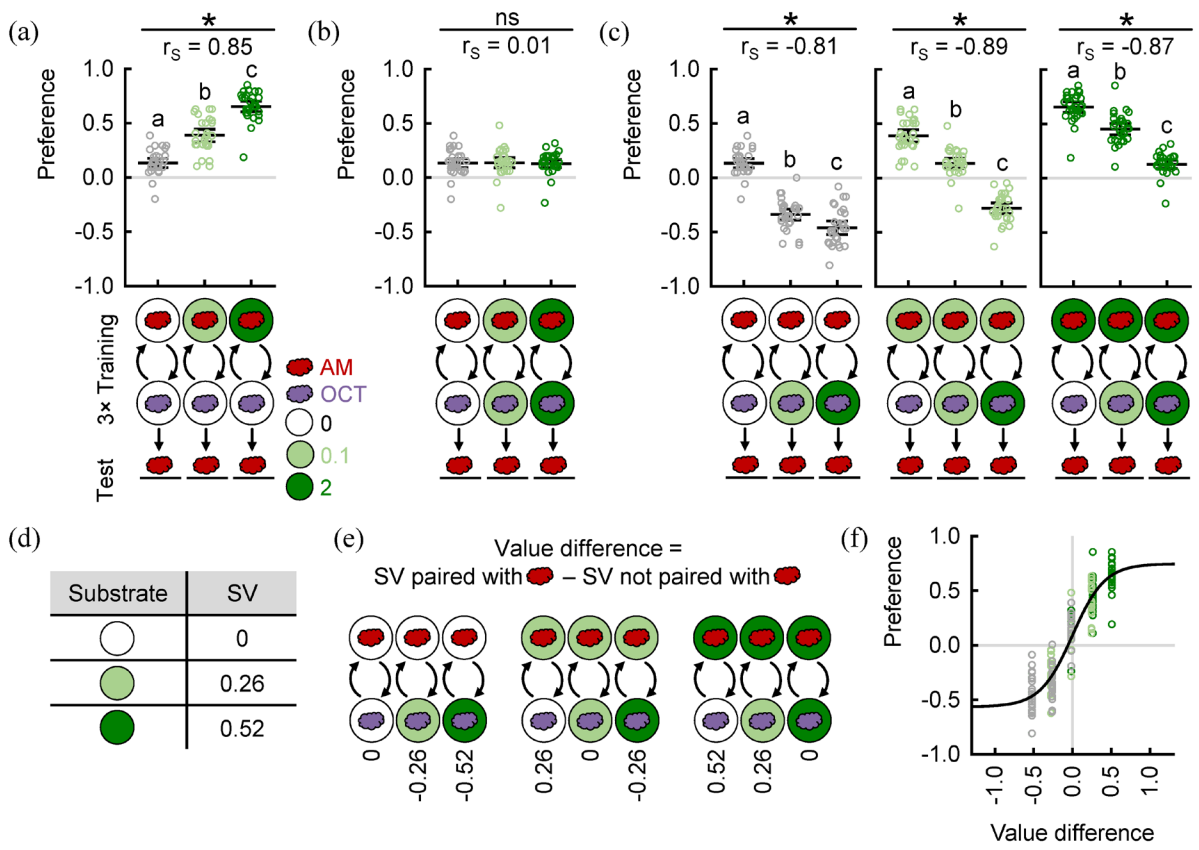

Figure 1. Larvae learn about the relative value of rewards. (a) Larvae were trained with AM paired with 0, 0.1 or 2 M fructose reward and OCT paired with no reward. (b) Larvae were trained with the same fructose concentration in both training dishes. (c) Larvae were trained with the fructose concentration paired with AM kept constant, but the concentration paired with OCT varied. (d) The subjective value (SV) of each reward was deduced from the animals' preference in (a). See Materials & Methods for details. (e) The (signed) difference in reward value for each experimental group. (f) AM preferences increased with increasing difference in reward value. A generalized logistic function was fitted to the data. Sample size is 30 each. Note that the groups displayed in (a-b) are also included in (c). * and "ns" indicate significant and not significant SRC tests, respectively, and r_S indicates the Spearman's correlation coefficient. Different small letters above each group indicate significant pairwise difference from each other (Mann-Whitney U tests). For all raw data along with the detailed results of the statistical tests, see Supplementary table S1.

be no sign of associative learning in this experiment, because both alternating substrates were the same and therefore the relative value of the reward was zero (neither better nor worse than the alternative). Indeed, we found that the animals all displayed the same baseline AM preference, irrespective of the reward strength (Fig. 1b). This suggests that they learned about the relative and not the absolute reward value.

230 Next, we modified the experiment such that AM was paired with no reward and thus the
231 absolute reward value was zero, but the alternative featured different reward concentrations (Fig.
232 1c, left). We found that the AM preference was decreasing with increasing reward strength that
233 had been paired with OCT. Again, this fits to the hypothesis of relative value learning: as the
234 substrate paired with AM had been less good than the alternative, its relative value was negative.
235 Finally, we repeated the same experiment, but paired AM either with the weak reward (Fig. 1c,
236 middle) or the strong reward (Fig. 1c, right). In each case, the AM preference declined along with
237 the reward strength that had been paired with OCT, and in each case the result indicated that the
238 animals' behaviour was determined by the relative reward value associated with AM. Strikingly,
239 this even caused the larvae to avoid AM after it had been paired with a reward, but the alternative
240 substrate had been even better (Fig. 1c, 6th group from the left). Thus, we concluded that *D.*
241 *melanogaster* larvae learn about the relative value of a reward, not the absolute value. We obtained
242 an analogous result when we repeated the complete experiment but tested for the OCT preference
243 instead of AM (Fig. S3a-c).
244 In this study we used both fructose reward and quinine punishment (see below). Quinine
245 memories have been repeatedly shown to be only retrieved when the animals are tested in presence
246 of quinine [39, 53, 70-72]. Therefore, and in order to ensure comparability throughout the study,
247 we used 5 mM quinine test Petri dishes for all experiments, including experiments that only used
248 reward during training. Previous work has demonstrated that the retrieval of fructose reward
249 memories is not influenced by the presence of quinine [53]. To further control whether the test
250 condition does not influence our results, we repeated the complete experiment once more, but with
251 the test performed on a pure, tasteless Petri dish, and obtained very similar results irrespective of
252 the test condition (Fig. S3d-f).
253 To illustrate how the difference in value between the two substrates determines the preference,
254 we deduced the subjective values of the two rewards directly from the animals' behaviour during
255 the experiment in Fig. 1a, as the average preference of a group that received AM-reward training
256 minus the average of the baseline preference group (Fig. 1d; for more details, see Materials &
257 Methods). Then, we determined the value difference in each of the experimental groups of Fig. 1c
258 by subtracting the subjective value of the substrate paired with OCT from the value of the substrate
259 paired with AM (Fig. 1e). Finally, we plotted the AM preferences of all groups over the value
260 differences and used a generalized logistic function to describe the relationship between them (Fig.

1f, Fig. S3c,f). Across the data of this study, a generalized logistic function fits the data better than a linear function (Fig. S4).

In summary, we found that larvae learned about the relative value of rewards, and that the value difference between the options during training determined the odour preference in a logistic manner.

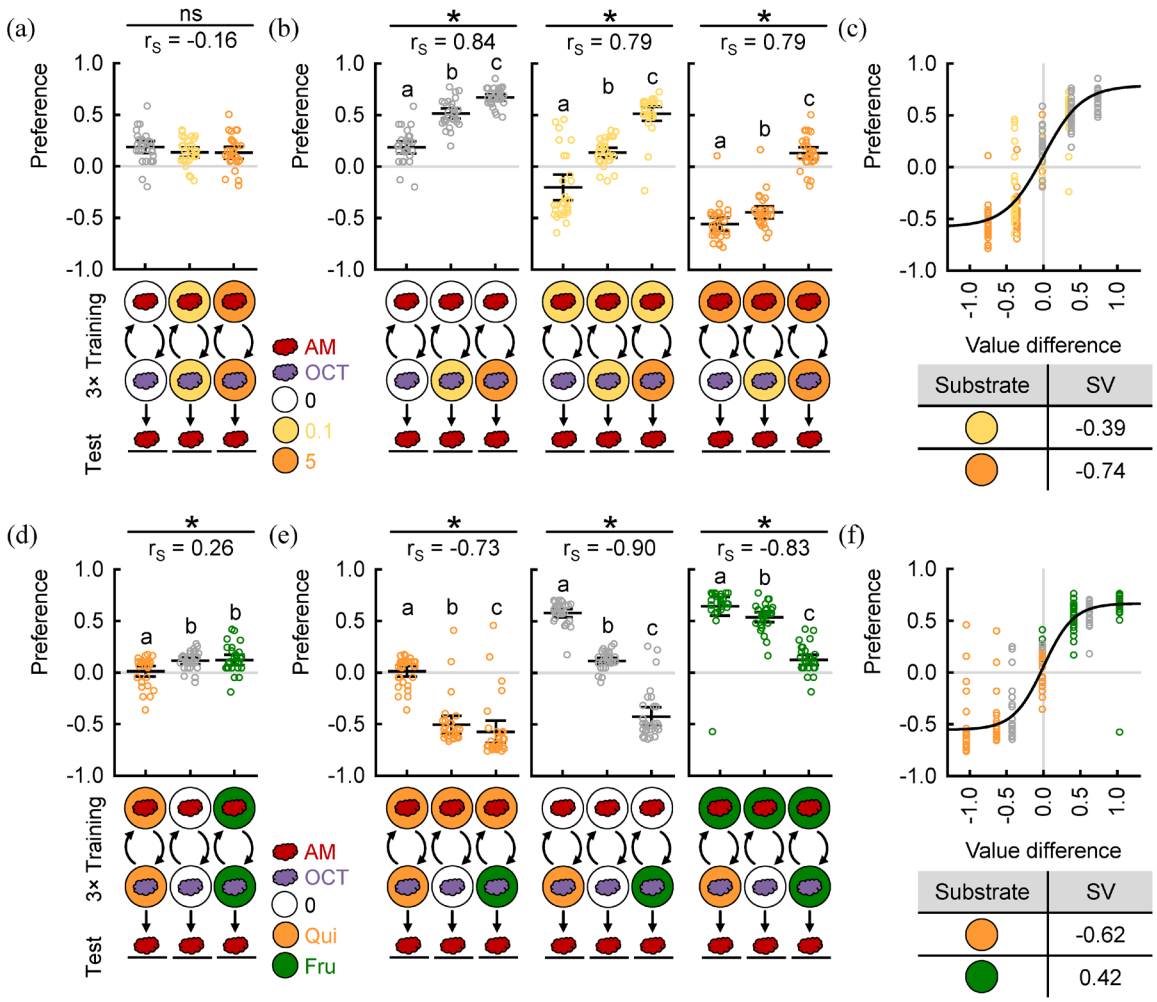

Figure 2. Larvae learn about the relative value of punishments. (a-c) Larvae were trained with two odours, AM and OCT, and different concentrations of quinine (0, 0.1 and 5 mM). (a) The same quinine concentration was presented in both training dishes. (b) The quinine concentration paired with AM was kept constant, but the concentration paired with OCT varied. (c) AM preferences increased with increasing difference in punishment value. The SV of each punishment is displayed below the plot. (d-f) Same as in (a-c), but with combinations of 5 mM quinine as punishment and 2 M fructose as reward. Sample size is 30 each. Note that the groups displayed in (a) and (d) are also included in (b) and (e), respectively. For further details, see Fig. 1. For all raw data along with the detailed results of the statistical tests, see Supplementary table S1.

Larvae learn about the relative value of punishments

Next, we asked whether the larvae learn about the absolute or relative value of punishments. To answer this, we performed a similar experiment as shown in Figure 1, but used two different concentrations of quinine as punishment, based on an initial experiment (Fig. S5a-c). We ensured that the larvae can still detect and process the lower of the two quinine concentrations immediately after being exposed to the higher concentration, although they showed a trend of a slightly reduced avoidance (Fig. S5d).

Similar to the previous findings, the animals all displayed the same baseline AM preference after they had been trained with the same quinine concentration in both training dishes, irrespective of the punishment strength (Fig. 2a). When AM was paired with no punishment and thus the absolute punishment value was zero, but the alternative featured different punishment concentrations (Fig. 2b, left), we found that the AM preference was increasing with increasing punishment strength that had been paired with OCT. Finally, when AM had been paired either with the weak punishment (Fig. 2b, middle) or the strong punishment (Fig. 2b, right), in each case the AM preference increased along with the punishment strength that had been paired with OCT. Thus, we concluded that larvae learn about the relative value of a punishment, not the absolute value (Fig. 2c).

Larvae learn about the relative values of combined reward and punishment

After finding that the larvae learned about the relative value of rewards and punishments in separate experiments, we wondered what would happen if we combined reward and punishment in the same experimental setup. When both odours had been paired with the same reward or punishment, we saw a slight but significant difference in AM preference, mostly due to a decreased AM preference after training with quinine (Fig. 2d). We note that this group is an exact repetition of the one in Fig. 2a (right-most group) in which we saw no reduced preference. This difference may point to the level of behavioural variability in our experiments and cautions against over-interpreting small correlations, even if they are statistically significant. When keeping the reward or punishment paired with AM constant, we found, once more, a clear change in AM preference along with the substrate that had been paired with OCT, fitting to the hypothesis of relative value learning (Fig. 2e). We note that pairing one odour with reward, the other one with punishment led

296 to stronger negative or positive preferences than training with only the reward or only the
297 punishment. Thus, we again concluded that larvae learn about the relative values of reward and
298 punishment, not the absolute values. In this experiment, as the value differences between the
299 substrates were particularly large, the logistic relationship (as compared to a simple linear one)
300 between the value difference and the odour preference was particularly striking (Fig. 2f).
301
302 *Larvae can learn about relative reward values with only one odour*
303 In addition to the differential, two-odour training paradigm, a one-odour version is used regularly
304 [62, 73-75]. We next asked whether the larvae would also learn the relative reward value in such
305 a paradigm, or whether the presence of two odours during the training was required. We repeated
306 the experiment of Fig. 1 but trained the larvae by pairing the various substrates with only one odour,
307 either AM (Fig. 3a-c) or OCT (Fig. 3d-f).
308 When we trained the larvae with the same fructose concentration in both training dishes, they
309 showed increased AM preferences with increased reward strength which indicates that they learned
310 about the absolute value of the reward (Fig. 3a). Next, we trained the animals such that AM was
311 paired with no reward and thus the absolute reward value was zero, but the alternative featured
312 different reward strengths (Fig. 3b, left). We observed that the AM preference was lower when the
313 alternative substrate had been fructose, suggesting that also the relative value had been learned.
314 When we paired AM with reward, we observed a similar tendency, but dampened the stronger the
315 reward paired with AM had been (Fig. 3b, middle, right).
316 We repeated the same experiment using OCT (Fig. 3d-f). Here, the OCT preference did not
317 change when the same substrate was used in both training dishes (Fig. 3d), and when the fructose
318 concentration paired with OCT was kept constant, the OCT preference was consistently decreasing
319 with increasing fructose concentrations not paired with OCT (Fig. 3e-f). These results very much
320 matched those from the two-odour experiments and clearly indicated that the larvae learned about
321 the relative but not the absolute value of the rewards.
322 Finally, we decided to try benzaldehyde (BA), another odour that is often used in larval learning
323 experiments [69, 76], and found results that closely matched those seen with OCT and indicated
324 relative value learning (Fig. 3g-i).
325

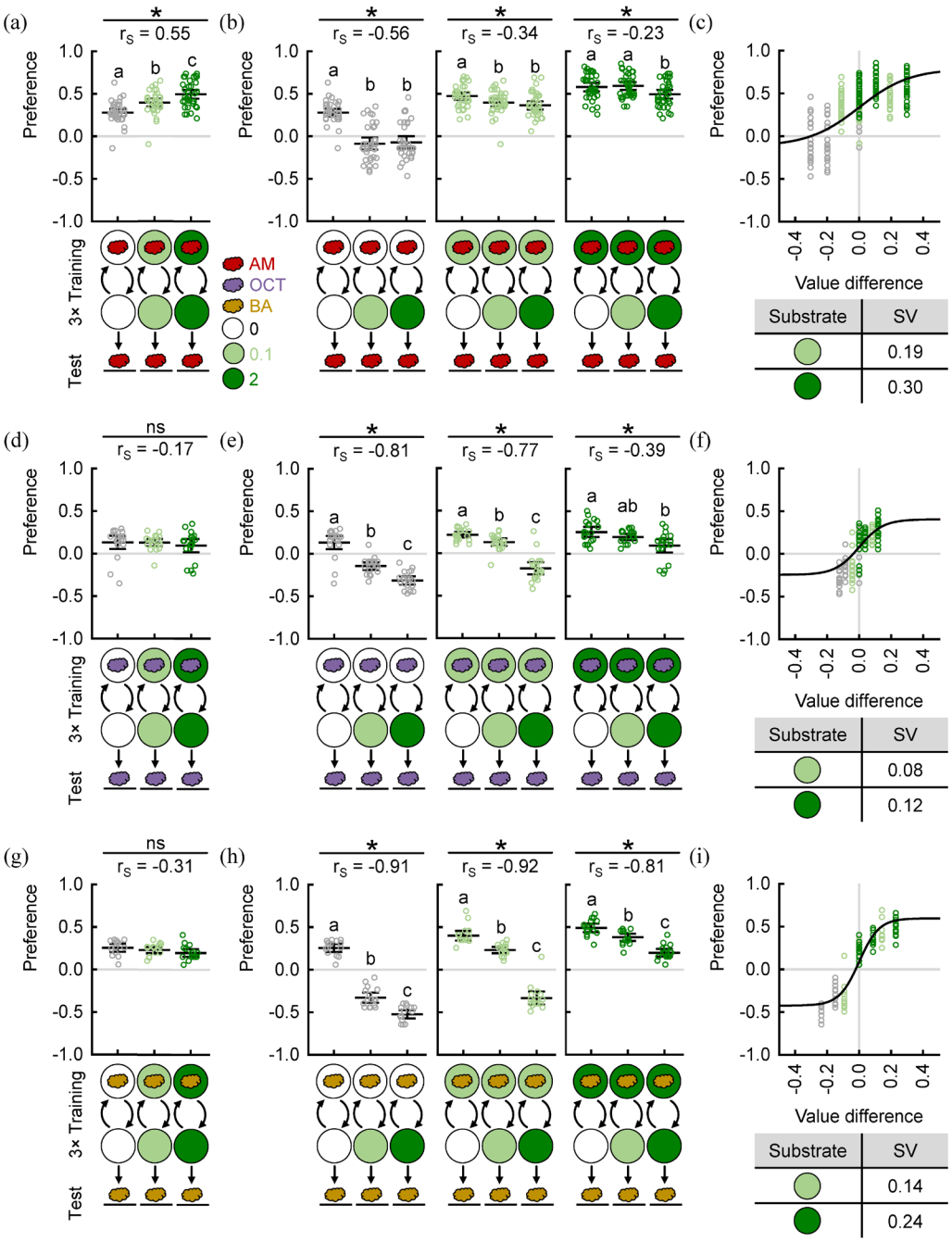

Figure 3. Larvae can learn about the relative value of rewards with one-odour training. (a-c) Larvae were trained only with AM. All other experimental details as in Fig. 1. (a) Larvae were trained with the same fructose concentration in both training dishes. (b) Larvae were trained with the fructose concentration paired with AM kept constant, but the concentration unpaired with AM varied. (c) AM preferences increased with increasing difference in reward value. The SV of each reward is displayed below the plot. (d-f) Same as in (a-c), but for OCT as odour. (g-i) Same as in (a-c), but for BA as odour. Sample size is (a-c) 36, (d-f) 20, (g-i) 16 each. Note that the groups displayed in (a), (d) and (g) are also included in (b), (e) and (h), respectively. For further details, see Fig. 1. For all raw data along with the detailed results of the statistical tests, see Supplementary table S1.

In summary, we found that the larvae can also learn about the relative reward value with only one odour during training. However, the extent to which the relative but not the absolute reward value was learned depended on the trained odour and the reward strength paired with it.

Larvae learn relatively independent of the training sequence

Finally, we asked whether relative value learned depended on the training sequence which determines the information available to the animal at each step: if AM was presented in the second training phase, the animals could have a memory about the reward strength of the previous phase, and associate the relative increase or decrease with AM. In contrast, if AM was presented first, this information would be not available at the time of AM presentation.

All experiments so far included three training cycles that followed each other immediately and thus do not allow for a precise control of the information available at each step (Fig. S1). Therefore, we switched to a one-trial training paradigm [77]. The training was performed as described for Fig. 1, but the larvae were exposed to each odour-substrate combination only once before they were tested for their AM preference. We first trained the animals such that OCT was always presented first and AM second (Fig. 4a-c). The results were very similar to the three-trial experiments and clearly indicated relative value learning. Then, we trained the animals with the other sequence, presenting AM first and OCT second (Fig. 4d-f). When both odours had been paired with the same reward strength, we found a slight but significant increase of the AM preference with increasing reward strength (Fig. 4d). Surprisingly, when we kept the substrate paired with AM constant, the AM preference decreased with increasing reward strength paired with OCT, indicating relative value learning (Fig. 4e). Overall, the results were very similar between both training sequences. As an additional control, we tested whether the exposure to sugar rewards alone may affect odour preferences in a way that can lead to the observed results (Fig. 4d-f). We found that exposing the animals to reward in absence of any odour decreased the subsequent preference for AM (Fig. S6a). In addition, the sequence of reward exposures mattered: increasing reward strengths from the first to the second dish decreased the subsequent AM preference, decreasing reward strengths increased it (Fig. S6b-c). This trend qualitatively matched the results of the previous experiment, but to a much smaller extent (S6c). Taken together, we concluded that the exposure to sugar alone cannot explain our odour-sugar training results and that, indeed, larvae can learn about the relative reward value irrespective of the training sequence.

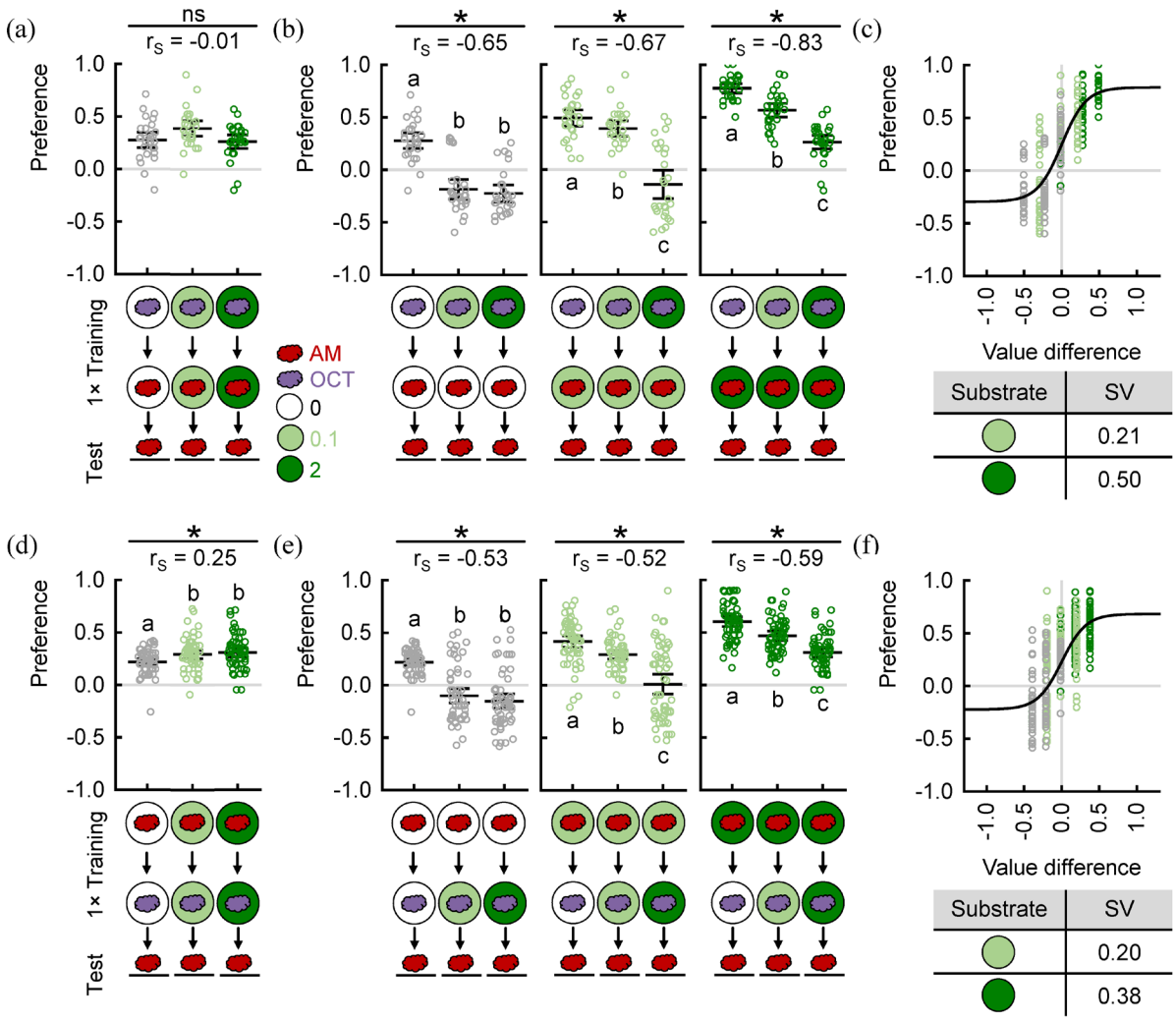

Figure 4. Larvae can learn the relative value of rewards independent of training sequence. Larvae were trained with two odours, AM and OCT, and different concentrations of fructose. Only a single training trial was performed, with (a-c) OCT or (d-f) AM being presented first. (a, d) The same fructose concentration was presented in both training dishes. (b, e) The fructose concentration paired with AM was kept constant, but the concentration paired with OCT varied. (c, f) AM preferences increased with increasing difference in reward value. The SV of each reward is displayed below the plot. Sample size is 28 each for (a-c) and 58 each for (d-f). Note that the groups displayed in (a) and (d) are also included in (b) and (e), respectively. For further details, see Fig. 1. For all raw data along with the detailed results of the statistical tests, see Supplementary table S1.

Discussion

In this study, we asked whether *D. melanogaster* larvae learn about the absolute or relative values of rewards and punishments during classical conditioning. We designed an experiment that allowed us to distinguish the two hypotheses by pairing each of two odours with rewards or punishments of different strength, followed by testing the preference of one odour alone to assess what the animals had associated with it. Our results clearly match the hypothesis that the values of both training situations were compared and the relative value was associated with one odour, both for rewards (Fig. 1) and punishments (Fig. 2), irrespective of which odour was tested (Fig. S3) as well as of the number and sequence of training trials (Fig. 4). Can these results explained by processes that do not require the comparison of the relative values?

Previous exposure of sugar rewards may change the animals' motivation and evaluation of further sugar rewards. Indeed, we observed that the larvae showed a reduced preference towards low-sugar when they had experienced the high concentration before (Fig. S2d). This may reduce the subjective reward value that is associated with an odour in the according learning condition and thus explain the differences between training regimen we observed. Alternatively, the learning processes in the two training phases may disrupt each other, a phenomenon called memory interference in Psychology [78-79]. Theoretically, the memory formed about sugar in the first training phase may inhibit learning about the odour-sugar association in the second phase [80]. Likewise, learning about the sugar in the second phase may impair the retrieval of the memory from the first phase [81]. Both processes could lead to differences between our training regimen. However, neither a reduced motivation nor impaired learning or retrieval capabilities can explain an inversion of memory valence – that is, that the animals form an aversive memory about the odour paired with the weaker of two rewards, as we observed it repeatedly (Fig. 1c, 3e, 3h, 4b, S3b, S3e; 6th group from the left).

In contrast, this result is in line with one of the experimental hallmarks of relative value learning in humans: that it can lead to objectively suboptimal decisions. Humans tend to choose options that had been better than alternatives in previous experiences over options that predict a higher absolute value [8, 14-15]. The larvae avoided an odour that had been paired with a potent reward if the alternative had been even better – although no better option was available at the moment of the choice test. This prevented the larvae to find an otherwise positively rated sugar source. Thus,

387 it seems that the way how brains evaluate rewards or punishments during learning is shared across
388 the animal kingdom from insect larvae to humans.
389 Notably, our results do not exclude that larvae can learn about the absolute values of rewards.
390 What animals learn depends strongly on the specific training procedure. We presented the two
391 training situations in an alternating fashion, each of them paired with a different odour – this
392 procedure likely facilitated a direct comparison of the reward values. A blocked design, or using
393 only one odour during the training, could potentially favour absolute value learning [12]. Indeed,
394 when training only with AM, we observed partially absolute, partially relative value learning (Fig.
395 3a-c). When using other odours, however, we found complete relative value learning, matching
396 the results of the differential, two-odour training (Fig. 3d-i). The reason for this discrepancy is not
397 known. In any case, this result shows that relative value learning does not require the presence of
398 two alternating stimuli that are associated with the different rewards.
399
400 *Relative value learning is consistent with previous studies*
401 To our knowledge, this is the first demonstration of relative value learning in *D. melanogaster*.
402 Previous studies in both larvae and adult flies had used different punishment strengths during
403 training, each paired with a given odour, and then had tested the animals' choice between these
404 odours [39-40, 44]. These studies demonstrated that the animals can learn how bad a given
405 punishment is, but they could not distinguish between relative and absolute value learning, as it
406 remains possible that the animals learned the absolute value paired with each odour and then
407 compared during the choice test which of the odours predicted the better outcome. The critical
408 advantage of our design is that we tested only a single odour and therefore could directly assess
409 what the animals had learned about it.
410 A number of studies used an experimental paradigm similar to ours with two odours during the
411 training but only one of them during the test [64, 82-84]. In addition, a one-odour training version
412 is widely used in larvae [49, 53, 61-62, 72-76, 85-86]. In these studies, a single reward or
413 punishment was paired with one odour, while the other training situation was neutral and therefore
414 had a value of zero. Their results are not conclusive for, yet consistent with the hypothesis of
415 relative value learning.
416 For example, one-odour experiments with rewards always consisted of two types of training
417 regimen: either the odour was paired with the reward while the other training situation featured no

odour on a neutral substrate (so-called paired training), or the odour was paired with the neutral substrate while the reward was presented without an odour (so-called unpaired training). For paired training, both absolute and relative value learning would lead to the same outcome, an increased odour preference compared to the baseline. For unpaired training, in contrast, absolute value learning would lead to a baseline level of odour preference because the odour had not been paired with any reward, whereas relative value learning would lead to a preference lower than the baseline because the odour had been paired with the less rewarding of the two substrates. The latter is precisely what had been found in several previous studies and was confirmed in this study (Fig. 3) [61-62, 64, 72]. Analogous results have been observed when using punishments instead of rewards [64, 72] as well when using a training with two odours (Fig. 1c, 2b) [64]. Using a different experimental design, similar results have been found in adults, too. Here, the resulting memories were called "CS+" and "CS-" memories [83].

These observations gave rise to the proposal of two separate learning processes after paired and unpaired training, leading to memories of opposite valences and with separate neurobiological mechanisms [62, 64, 72, 83, 87-88]. However, relative value learning provides a unified framework to explain the results of both kinds of training: the memory's valence depends on whether the trained odour had been paired with the better or worse of the training situations. This new perspective should be incorporated into the current theoretical, computational and neuronal working models of associative learning in *D. melanogaster*. In the following, we will briefly discuss one plausible neuronal working model.

A potential neuronal circuit for relative value learning

In *D. melanogaster*, olfactory associative memories are established in the mushroom body, a major insect memory centre (Fig. 5a) [38, 89]. In brief, odours and other stimuli are coded by specific combinations of Kenyon cells (KCs), the intrinsic neurons of the mushroom body. The KCs are intersected by (mostly) dopaminergic modulatory input neurons (DANs) and mushroom body output neurons (MBONs). The innervations of DANs and MBONs compartmentalize the mushroom body into microcircuits between DANs, KCs and MBONs that act as functional units. As a rule, DANs conveying reward signals pair up with avoidance-promoting MBONs, and DANs conveying punishment signals pair up with approach-promoting MBONs. DAN signals are received by the axonal terminals of all KCs and modulate the KC-MBON synapses. The synaptic

weights of the KC-MBON synapses thus contain the associative strengths of all stimuli coded by the KCs. Crucially, studies in adults found that the same DAN signal can have opposite effects on internal KC signalling when a KC is active versus inactive [90], and that the KC-MBON synapses can be depressed or potentiated depending on the training contingency [91].

The MBONs gather input from all KCs and on the one hand convey this information towards motor control, tilting the behaviour of the animals towards approach or avoidance, depending on the memory [86, 92-94]. On the other hand, they feed this information back either directly, or via one and two step feedback relays to the DANs [74]. The DANs that receive the feedback may innervate the same or other compartments and may even code for the opposite valence [74].

The predominant working model for associative learning in *D. melanogaster* proposes that when a DAN reward signal is received by a KC while it is active (i.e. while the respective odour is present), its output synapses are depressed to an extent matching the reward strength (Fig. 5b, left). This reduces the activity of the avoidance-promoting MBON while leaving the activity of approach-promoting MBONs unchanged and thus shifts the mushroom body's net output towards approach [86, 92-94]. This process can be seen as the basis for absolute value learning: an odour is associated with the reward strength it coincides with [44].

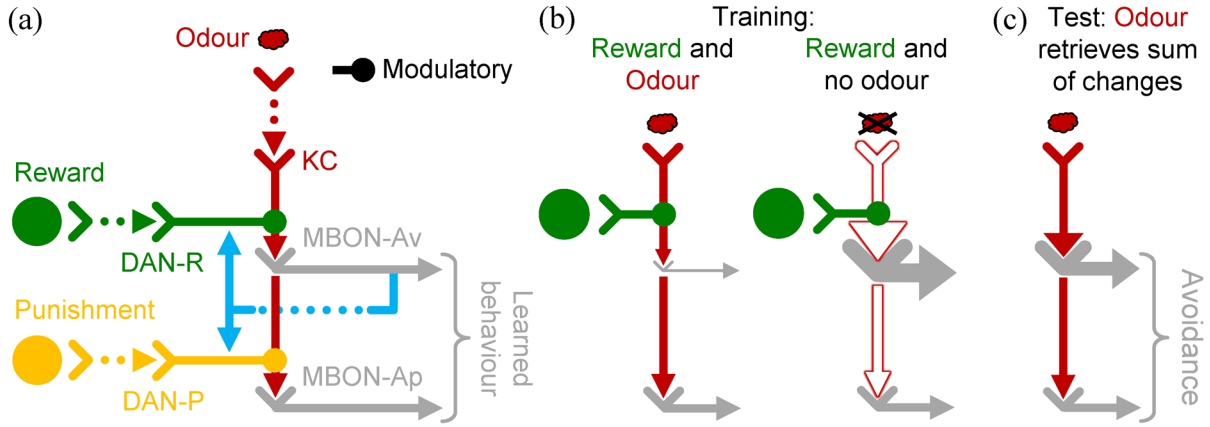

Figure 5. Simplified potential circuitry underlying relative value learning. (a) Mushroom body circuit for associative learning. DAN: dopaminergic neuron; KC: Kenyon cell; MBON: mushroom body output neuron. Two sample compartments are shown, one receiving reward, the other one punishment input. The feedback from MBONs to DANs (blue, only displayed for one compartment) can be direct or indirect and excitatory or inhibitory. Based on [49, 74, 85]. (b) Simple working model of absolute and relative value learning. Left: coincidence of odour and reward depresses the synapses of odour-activated KCs and an avoidance-promoting MBON. Right: a reward received by an inactive KC potentiates the KC-MBON synapses. (c) The KC-MBON synapses integrate the modulations by previous rewards in presence and absence of the odour. In the shown example, the net effect is a potentiated synapse and thus odour avoidance.

465 For relative value learning, the crucial question is how the reward strengths can be compared
466 and the relative value can be computed. In mammals, DANs have been found to compute an
467 associative prediction error that can be interpreted as the difference between the current and
468 previously learned reward strengths and thus as a computation of relative reward value [95-97].
469 Many computational models assume a similar mechanism in *D. melanogaster*, driven via the
470 feedback from the MBONs to the DANs, but hard evidence for such a mechanism is lacking [74,
471 98-99].
472 We propose that the computation of the relative value can occur in the KCs themselves: when
473 a DAN signal is received by an inactive KC (i.e. the respective odour is not present), its output
474 synapses are potentiated to an extent matching the reward strength, leading to a shift in the
475 mushroom body's net output towards avoidance (Fig. 5b, right) [44, 64, 91]. If rewards of different
476 strengths are presented in presence and absence of an odour, like in our experiment, the final
477 strength of the KC-MBON synapses reflects an integration, in the simplest case a sum, of all
478 modulations during training (Fig. 5c). In other words, the relative reward value is computed within
479 the KC-MBON synapse, as an integration of all positive and negative molecular modulations of
480 the synapse strength.
481 Interestingly, the proposed model predicts that KC-MBON synapses get potentiated whenever
482 a reward is encountered while a KC is not active. This should decrease the preference for any
483 odour that was not present at that moment. Indeed, we found that exposing the animals to sugar
484 reward alone decreases preferences to a novel odour, although the effect was quite small (Fig. S6a).
485 We also found that the presence of a second training odour was not crucial for relative value
486 learning – what mattered was the difference in reward strength in presence and absence of the to-
487 be-tested odour (Fig. 1, Fig. 3).
488 Although the proposed working model is very simple, it can be easily expanded to account for
489 additional considerations. For example, humans have been found to compare reward strengths
490 within each learning session [8], and mammals often learn in a context-dependent manner [100-
491 102]. Possibly, also the KCs that get potentiated are limited, e.g. to those that code for stimuli that
492 had been encountered recently, or in the same experimental context. Such a mechanism would
493 prevent an unnecessary potentiation of all KC-MBON synapses whenever a reward is encountered.
494 In adult flies, reward-signalling DANs have been found to be required for memories about the
495 punishment-unpaired odour, but not about the punishment-paired odour [83]. Another study found

a MBON pathway that encodes a “better-than-before” signal and activates a reward-signalling dopaminergic pathway, leading to a secondary reward memory in a different mushroom body compartment [44]. In other words, when an odour was paired with the weaker of two rewards or punishments, the modulations of the KC-MBON synapses we propose here may, as a second step, establish a memory of the opposite valence via the opposite-valence DANs. Future research will elucidate these and other pending questions regarding the detailed mechanism of how brains integrate the relative values of options into associative learning, in larvae as well as in other animals. The *D. melanogaster* larva, with its small brain and excellent tools for transgenic manipulations, is perfectly suited to contribute to this endeavour.

Acknowledgements

We thank A. Fukushima, B. Gerber, G. Jocham, M. Mizunami, P. Tobler and M.L.M. Wang for fruitful discussions and N. Toshima for helpful comments on the manuscript. We are grateful to the Ministry of Education, Culture, Sports and Technology (MEXT), Government of Japan, for financial support to S. Rahman through the MEXT Scholarship.

References

[1] Fellows LK. 2018. The neuroscience of human decision-making through the lens of learning and memory. *Curr. Top. Behav. Neurosci.* **37**:231–251. (doi:10.1007/7854_2016_468).

[2] St-Amand D, Sheldon S, Otto AR. 2018. Modulating episodic memory alters risk preference during decision-making. *J. Cogn. Neurosci*. **30**:1433–1441. (doi:10.1162/jocn_a_01253).

[3] Biderman N, Bakkour A, Shohamy D. 2020. What are memories for? The hippocampus bridges past experience with future decisions. *Trends Cogn. Sci*. **24**:542–556. (doi:10.1016/j.tics.2020.04.004).

[4] Seymour B, McClure SM. 2008. Anchors, scales and the relative coding of value in the brain. *Curr. Opin. Neurobiol*. **18**:173–178. (doi:10.1016/j.conb.2008.07.010).

[5] Vlaev I, Chater N, Stewart N, Brown GD. 2011. Does the brain calculate value? *Trends Cogn. Sci.* **15**:546–554. (doi:10.1016/j.tics.2011.09.008).

[6] Rangel A, Clithero JA. 2012. Value normalization in decision making: theory and evidence. *Curr. Opin. Neurobiol*. **22**:970–981. (doi:10.1016/j.conb.2012.07.011).

[7] Palminteri S, Khamassi M, Joffily M, Coricelli G. 2015. Contextual modulation of value signals in reward and punishment learning. *Nat. Commun*. **6**:9096. (doi:10.1038/ncomms9096).

[8] Klein TA, Ullsperger M, Jocham G. 2017. Learning relative values in the striatum induces violations of normative decision making. *Nat. Commun*. **8**:16033. (doi:10.1038/ncomms16033).

[9] Palminteri S, Lebreton M. 2021. Context-dependent outcome encoding in human reinforcement learning. *Curr. Opin. Behav. Sci.* **41**:144–151. (doi:10.1016/j.cobeha.2021.06.006).

[10] Juechems K, Altun T, Hira R, Jarvstad A. 2022. Human value learning and representation reflect rational adaptation to task demands. *Nat. Hum. Behav*. **6**:1268–1279. (doi:10.1038/s41562-022-01360-4).

[11] Hayes WM, Wedell DH. 2023. Reinforcement learning in and out of context: The effects of attentional focus. *J. Exp. Psychol. Learn. Mem. Cogn*. **49**:1193–1217. (doi:10.1037/xlm0001145).

[12] Hayes WM, Wedell DH. 2023. Effects of blocked versus interleaved training on relative value learning. *Psychon. Bull. Rev*. **30**:1895–1907. (doi:10.3758/s13423-023-02290-6).

[13] Bavard S, Palminteri S. 2023. The functional form of value normalization in human reinforcement learning. *Elife*. **12**:e83891. (doi:10.7554/eLife.83891).

[14] Bavard S, Lebreton M, Khamassi M, Coricelli G, Palminteri S. 2018. Reference-point centering and range-adaptation enhance human reinforcement learning at the cost of irrational preferences. *Nat. Commun*. **9**:4503. (doi:10.1038/s41467-018-06781-2).

[15] Bavard S, Rustichini A, Palminteri S. 2021. Two sides of the same coin: beneficial and detrimental consequences of range adaptation in human reinforcement learning. *Sci. Adv*. **7**:eabe0340. (doi:10.1126/sciadv.abe0340).

[16] Louie K, Glimcher PW. 2012. Efficient coding and the neural representation of value. *Ann. N.Y. Acad. Sci.* **1251**:13-32. (doi:10.1111/j.1749-6632.2012.06496.x).

[17] Bhui R, Gershman SJ. 2018. Decision by sampling implements efficient coding of psychoeconomic functions. *Psychol. Rev.* **125**:985–1001. (doi:10.1037/rev0000123).

[18] Burke JC, Baddeley M, Tobler PN, Schultz W. 2016. Partial Adaptation of Obtained and Observed Value Signals Preserves Information about Gains and Losses. *J. Neurosci.* **36**:10016-10025. (doi:10.1523/JNEUROSCI.0487-16.2016).

[19] Stephens DW, Krebs JR. 1986. Foraging Theory. Princeton University Press.

[20] Dukas R. 1998. Cognitive Ecology: The Evolutionary Ecology of Information Processing and Decision Making. University of Chicago Press.
[21] Charnov EL.1976. Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.* **9**:129-136. (doi:10.1016/0040-5809(76)90040-X).
[22] McNamara JM. 1982. Optimal patch use in a stochastic environment. *Theor. Popul. Biol.* **21**:269-288. (doi:10.1016/0040-5809(82)90018-1).
[23] McNamara JM, Houston AI. 1986. The common currency for behavioral decisions. *Am. Nat.* **127**:358-378. (doi:10.1086/284489).
[24] Real LA. 1991. Animal choice behavior and the evolution of cognitive architecture. *Science* **253**:980-986. (doi:10.1126/science.1887231).
[25] McNamara JM, Fawcett TW, Houston AI. 2013. An adaptive response to uncertainty generates positive and negative contrast effects. *Science.* **340**:1084–1086. (doi:10.1126/science.1230599).
[26] Evans MR et al. 2013. Do simple models lead to generality in ecology? *Trends Ecol. Evol.* **28**:578-583. (doi: 10.1016/j.tree.2013.05.022).
[27] Fawcett TW, Fallenstein B, Higginson AD, Houston AI, Mallpress DEW, Trimmer PC, McNamara JM. 2014. The evolution of decision rules in complex environments. *Trends Cogn Sci.* **18**:153–161. (doi:10.1016/j.tics.2013.12.012).
[28] Hemingway CT, Pimplikar S, Muth F. 2024. Wild bumblebees use both absolute and relative evaluation when foraging. *Behav. Ecol.* **35**:arad112. (doi:10.1093/beheco/arad112). [29] Belke TW. 1992. Stimulus preference and the transitivity of preference. *Anim. Learn. Behav*. **20**:401–406. (doi:10.3758/BF03197963).
[30] Flaherty CF, Turovsky J, Krauss KL. 1994. Relative hedonic value modulates anticipatory contrast. *Physiol. Behav*. **55**:1047–1054. (doi:10.1016/0031-9384(94)90386-7).
[31] Waite TA. 2001. Background context and decision making in hoarding gray jays. *Behav. Ecol*. **12**:318–324. (doi:10.1093/beheco/12.3.318).
[32] Tobler PN, Fiorillo CD, Schultz W. 2005. Adaptive coding of reward value by dopamine neurons. *Science*. **307**:1642–1645. (doi:10.1126/science.1105370).
[33] Pompilio L, Kacelnik A. 2010. Context-dependent utility overrides absolute memory as a determinant of choice. *Proc. Natl Acad. Sci. USA*. **107**:508–512. (doi:10.1073/pnas.0907250107).
[34] Freidin E, Kacelnik A. 2011. Rational choice, context dependence, and the value of information in European starlings (Sturnus vulgaris). *Science.* **334**:1000–1002. (doi:10.1126/science.1209626).
[35] Webber ES, Chambers NE, Kostek JA, Mankin DE, Cromwell HC. 2015. Relative reward effects on operant behavior: incentive contrast, induction and variety effects. *Behav. Process*. **116**:87–99. (doi:10.1016/j.beproc.2015.05.003).
[36] Giurfa M. 2015. Learning and cognition in insects. *Wiley Interdiscip. Rev. Cogn. Sci.* **6**:383-395. (doi: 10.1002/wcs.1348).
[37] Mizunami M. 2021. What Is Learned in Pavlovian Conditioning in Crickets? Revisiting the S-S and S-R Learning Theories. *Front. Behav. Neurosci.* **15**:661225. (doi: 10.3389/fnbeh.2021.661225).
[38] Modi MN, Shuai Y, Turner GC. 2020. The *Drosophila* mushroom body: from architecture to algorithm in a learning circuit. *Annu. Rev. Neurosci*. **43**:465–484. (doi:10.1146/annurev-neuro-080317-062133).

[39] Gerber B, Hendel T. 2006. Outcome expectations drive learned behaviour in larval *Drosophila*. *Proc. Biol. Sci*. **273**:2965–2968. (doi:10.1098/rspb.2006.3673).
[40] Yin Y, Chen N, Zhang S, Guo A. 2009. Choice strategies in *Drosophila* are based on competition between olfactory memories. *Eur. J. Neurosci*. **30**:279–288. (doi:10.1111/j.1460-9568.2009.06821.x).
[41] Riveros AJ, Gronenberg W. 2012. Decision-making and associative color learning in harnessed bumblebees (*Bombus impatiens*). *Anim. Cogn*. **15**:1183–1193. (doi:10.1007/s10071-012-0542-6).
[42] Price RI, Grüter C, Hughes WO, Evison SE. 2016. Symmetry breaking in mass-recruiting ants: extent of foraging biases depends on resource quality. *Behav. Ecol. Sociobiol*. **70**:1813–1820. (doi:10.1007/s00265-016-2187-y).
[43] Nityananda V, Chittka L. 2021. Different effects of reward value and saliency during bumblebee visual search for multiple rewarding targets. *Anim. Cogn.* **24**:803–814. (doi:10.1007/s10071-021-01479-3).
[44] Villar ME, Pavão-Delgado M, Amigo M, Jacob PF, Merabet N, Pinot A, Perry SA, Waddell S, Perisse E. 2022. Differential coding of absolute and relative aversive value in the *Drosophila* brain. *Curr. Biol.* **32**:4576–4592.e5. (doi:10.1016/j.cub.2022.08.058).
[45] Pompilio L, Kacelnik A, Behmer ST. 2006. State-dependent learned valuation drives choice in an invertebrate. *Science*. **311**:1613–1615. (doi:10.1126/science.1123924).
[46] Gil M, De Marco RJ, Menzel R. 2007. Learning reward expectations in honeybees. *Learn. Mem.* **14**:491–496. (doi:10.1101/lm.618907).
[47] Sokolowski MB. 2001. *Drosophila*: genetics meets behaviour. *Nat. Rev. Genet.* **2**:879–890. (doi:10.1038/35098592).
[48] Schroll C, Riemensperger T, Bucher D, Ehmer J, Völler T, Erbguth K, Gerber B, Hendel T, Nagel G, Buchner E, Fiala A. 2006. Light-induced activation of distinct modulatory neurons triggers appetitive or aversive learning in *Drosophila* larvae. *Curr. Biol*. **16**:1741–1747. (doi:10.1016/j.cub.2006.07.023).
[49] Eichler K, Li F, Litwin-Kumar A, Park Y, Andrade I, Schneider-Mizell CM, Saumweber T, Huser A, Eschbach C, Gerber B, Fetter RD, Truman JW, Priebe CE, Abbott LF, Thum AS, Zlatic M, Cardona A. 2017. The complete connectome of a learning and memory centre in an insect brain. *Nat.* **548**:175–182. (doi:10.1038/nature23455).
[50] Michels B, Zwaka H, Bartels R, Lushchak O, Franke K, Endres T, Fendt M, Song I, Bakr M, Budragchaa T, Westermann B, Mishra D, Eschbach C, Schreyer S, Lingnau A, Vahl C, Hilker M, Menzel R, Kähne T, Leßmann V, Gerber B. 2018. Memory enhancement by ferulic acid ester across species. *Sci. Adv*. **4**:eaat6994. (doi:10.1126/sciadv.aat6994).
[51] Winding M, Pedigo BD, Barnes CL, Patsolic HG, Park Y, Kazimiers T, Fushiki A, Andrade IV, Khandelwal A, Valdes-Aleman J, Li F, Randel N, Barsotti E, Correia A, Fetter RD, Hartenstein V, Priebe CE, Vogelstein JT, Cardona A, Zlatic M. 2023. The connectome of an insect brain. *Science*. **379**:eadd9330. (doi:10.1126/science.add9330).
[52] Scherer S, Stocker RF, Gerber B. 2003. Olfactory learning in individually assayed *Drosophila* larvae. *Learn. Mem*. **10**:217–225. (doi:10.1101/lm.57903).
[53] Schleyer M, Saumweber T, Nahrendorf W, Fischer B, von Alpen D, Pauls D, Thum A, Gerber B. 2011. A behavior-based circuit model of how outcome expectations organize learned behavior in larval *Drosophila*. *Learn. Mem*. **18**:639–653. (doi:10.1101/lm.2163411).
[54] Ohyama T, Schneider-Mizell CM, Fetter RD, Aleman JV, Franconville R, Rivera-Alba M, Mensh BD, Branson KM, Simpson JH, Truman JW, Cardona A, Zlatic M. 2015. A multilevel

multimodal circuit enhances action selection in *Drosophila*. *Nature.* **520**:633–639. (doi:10.1038/nature14297).

[55] Lihoreau M, Clarke IM, Buhl C, Sumpter DJ, Simpson SJ. 2016. Collective selection of food patches in *Drosophila*. *J. Exp. Biol*. **219**:668–675. (doi:10.1242/jeb.127431).

[56] Gepner R, Wolk J, Wadekar DS, Dvali S, Gershow M. 2018. Variance adaptation in navigational decision making. *Elife*. **7**:e37945. (doi:10.7554/eLife.37945).

[57] Mancini N, Hranova S, Weber J, Weiglein A, Schleyer M, Weber D, Thum AS, Gerber B. 2019. Reversal learning in *Drosophila* larvae. *Learn. Mem.* **26**:424–435. (doi:10.1101/lm.049510.119).

[58] Rahman S, Modak C, Akter M, Alam MS. 2020. Role of MicroRNA genes miR-1000 and miR-375 in forming olfactory conditional memory in *Drosophila melanogaster*. *MicroRNA*. **9**:283–288. (doi:10.2174/2211536609666200204113403).

[59] Eschment M, Franz HR, Güllü N, Hölscher LG, Huh KE, Widmann A. 2020. Insulin signaling represents a gating mechanism between different memory phases in *Drosophila* larvae. *PLoS Genet.* **16**:e1009064. (doi:10.1371/journal.pgen.1009064).

[60] Croteau-Chonka EC, Clayton MS, Venkatasubramanian L, Harris SN, Jones BMW, Narayan L, Winding M, Masson JB, Zlatic M, Klein KT. 2022. High-throughput automated methods for classical and operant conditioning of *Drosophila* larvae. *Elife*. **11**:e70015. (doi:10.7554/eLife.70015).

[61] Sen E, El-Keredy A, Jacob N, Mancini N, Asnaz G, Widmann A, Gerber B, Thoener J. 2024. Cognitive limits of larval *Drosophila*: testing for conditioned inhibition, sensory preconditioning, and second-order conditioning. *Learn. Mem*. **31**:a053726. (doi:10.1101/lm.053726.122).

[62] Saumweber T, Husse J, Gerber B. 2011. Innate attractiveness and associative learnability of odors can be dissociated in larval *Drosophila*. *Chem. Senses.* **36**:223–235. (doi:10.1093/chemse/bjq128).

[63] Michels B, Saumweber T, Biernacki R, Thum J, Glasgow RDV, Schleyer M, Chen YC, Eschbach C, Stocker RF, Toshima N, Tanimura T, Louis M, Arias-Gil G, Marescotti M, Benfenati F, Gerber B. 2017. Pavlovian conditioning of larval *Drosophila*: an illustrated, multilingual, hands-on manual for odor-taste associative learning in maggots. *Front. Behav. Neurosci*. **11**:45. (doi:10.3389/fnbeh.2017.00045).

[64] Schleyer M, Fendt M, Schuller S, Gerber B. 2018. Associative learning of stimuli paired and unpaired with reinforcement: evaluating evidence from maggots, flies, bees, and rats. *Front. Psychol.* **9**:1494. (doi:10.3389/fpsyg.2018.01494).

[65] Holm S. 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Stat*. **6**:65–70. (http://www.jstor.org/stable/4615733).

[66] Richards FJ. 1959. A flexible growth function for empirical use. *J. Exp. Bot.* **10**:290–300. (http://www.jstor.org/stable/23686557).

[67] R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (http://www.R-project.org).

[68] Akaike H. 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **19**: 716-723 (doi: 10.1109/TAC.1974.1100705).

[69] Weber D, Richter V, Rohwedder A, Großjohann A, Thum AS. 2023. Learning and memory in *Drosophila* larvae. *Cold Spring Harb. Protoc*. **2023**:pdb.top107863. (doi:10.1101/pdb.top107863).

[70] El-Keredy A, Schleyer M, König C, Ekim A, Gerber B. 2012. Behavioural analyses of quinine processing in choice, feeding and learning of larval *Drosophil*a. *PLoS One.* **7**:e40525. (doi:10.1371/journal.pone.0040525).
[71] Schleyer M, Miura D, Tanimura T, Gerber B. 2015. Learning the specific quality of taste reinforcement in larval *Drosophila*. *Elife*. **4**:e04711. (doi:10.7554/eLife.04711).
[72] Paisios E, Rjosk A, Pamir E, Schleyer M. 2017. Common microbehavioral "footprint" of two distinct classes of conditioned aversion. *Learn. Mem*. **24**:191–198. (doi:10.1101/lm.045062.117).
[73] Schleyer M, Weiglein A, Thoener J, Strauch M, Hartenstein V, Kantar Weigelt M, Schuller S, Saumweber T, Eichler K, Rohwedder A, Merhof D, Zlatic M, Thum AS, Gerber B. 2020. Identification of dopaminergic neurons that can both establish associative memory and acutely terminate its behavioral expression. *J. Neurosci*. **40**:5990–6006. (doi:10.1523/JNEUROSCI.0290-20.2020).
[74] Eschbach C, Fushiki A, Winding M, Schneider-Mizell CM, Shao M, Arruda R, Eichler K, Valdes-Aleman J, Ohyama T, Thum AS, Gerber B, Fetter RD, Truman JW, Litwin-Kumar A, Cardona A, Zlatic M. 2020. Recurrent architecture for adaptive regulation of learning in the insect brain. *Nat. Neurosci*. **23**:544–555. (doi:10.1038/s41593-020-0607-9).
[75] Toshima N, Behrad A, Behnke F, Kaushik G, Weiglein A, Strauch M, Thoener J, Kobler O, Wang MLM, Dörr M, Schleyer M. 2025. Individual dopaminergic neurons induce unique, yet overlapping combinations of behavioural modulations including safety learning, memory retrieval and acute locomotion. *Elife*. **14**:RP106148. (doi:10.7554/eLife.106148.1).
[76] Widmann A, Artinger M, Biesinger L, Boepple K, Peters C, Schlechter J, Selcho M, Thum AS. 2016. Genetic dissection of aversive associative olfactory learning and memory in *Drosophila* larvae. *PLoS Genet.* **12**:e1006378. (doi:10.1371/journal.pgen.1006378).
[77] Weiglein A, Gerstner F, Mancini N, Schleyer M, Gerber B. 2019. One-trial learning in larval *Drosophila*. *Learn. Mem.* **26**:109–120. (doi:10.1101/lm.049106.118).
[78] Hardt O, Nader K, Nadel L. 2013. Decay happens: the role of active forgetting in memory, *Trends Cogn. Sci.* **17**:111-120. (doi:10.1016/j.tics.2013.01.001).
[79] de Oliveira Alvares L, Do-Monte FH. 2021. Understanding the dynamic and destiny of memories. *Neurosci. Biobehav. Rev.* **125**:592-607. (doi:10.1016/j.neubiorev.2021.03.009).
[80] Jacob PF, Vargas-Gutierrez P, Okray Z, Vietti-Michelina S, Felsenberg J, Waddell S. 2021. Prior experience conditionally inhibits the expression of new learning in Drosophila. *Curr. Biol.* **31**:3490-3503. (doi: 10.1016/j.cub.2021.05.056).
[81] Felsenberg J, Barnstedt O, Cognigni P, Lin S, Waddell S. 2017. Re-evaluation of learned information in *Drosophila*. *Nat.* **544**:240–244. (doi:10.1038/nature21716).
[82] Galili DS, Dylla KV, Lüdke A, Friedrich AB, Yamagata N, Wong JY, Ho CH, Szyszka P, Tanimoto H. 2014. Converging circuits mediate temperature and shock aversive olfactory conditioning in *Drosophila*. *Curr. Biol*. **24**:1712–1722. (doi:10.1016/j.cub.2014.06.062).
[83] Jacob PF, Waddell S. 2020. Spaced training forms complementary long-term memories of opposite valence in *Drosophila*. *Neuron.* **106**:977–991.e4. (doi:10.1016/j.neuron.2020.03.013).
[84] Zhao B, Sun J, Li Q, Zhong Y. 2021. Differential conditioning produces merged long-term memory in *Drosophila*. *Elife*. **10**:e66499. (doi:10.7554/eLife.66499).
[85] Saumweber T, Rohwedder A, Schleyer M, Eichler K, Chen YC, Aso Y, Cardona A, Eschbach C, Kobler O, Voigt A, Durairaja A, Mancini N, Zlatic M, Truman JW, Thum AS, Gerber B. 2018. Functional architecture of reward learning in mushroom body extrinsic neurons of larval *Drosophila*. *Nat. Commun.* **9**:1104. (doi:10.1038/s41467-018-03130-1).

[86] Eschbach C, Fushiki A, Winding M, Afonso B, Andrade IV, Cocanougher BT, Eichler K, Gepner R, Si G, Valdes-Aleman J, Fetter RD, Gershow M, Jefferis GS, Samuel AD, Truman JW, Cardona A, Zlatic M. 2021. Circuits for integrating learned and innate valences in the insect brain. *Elife*. **10**:e62567. (doi:10.7554/eLife.62567).

[87] Zhao C, Widmer YF, Diegelmann S, Petrovici MA, Sprecher SG, Senn W. 2021. Predictive olfactory learning in *Drosophila*. *Sci. Rep*. **11**:6795. (doi:10.1038/s41598-021-85841-y).

[88] Thoener J, Weiglein A, Gerber B, Schleyer M. 2022. Optogenetically induced reward and 'frustration' memory in larval *Drosophila melanogaster*. *J. Exp. Biol.* **225**:jeb244565. (doi:10.1242/jeb.244565).

[89] Thum AS, Gerber B. 2019. Connectomics and function of a memory network: the mushroom body of larval *Drosophila*. *Curr. Opin. Neurobiol.* **54**:146–154. (doi:10.1016/j.conb.2018.10.007).

[90] Handler A, Graham TGW, Cohn R, Morantte I, Siliciano AF, Zeng J, Li Y, Ruta V. 2019. Distinct dopamine receptor pathways underlie the temporal sensitivity of associative learning. *Cell*. **178**:60–75.e19. (doi:10.1016/j.cell.2019.05.040).

[91] Cohn R, Morantte I, Ruta V. 2015. Coordinated and compartmentalized neuromodulation shapes sensory processing in *Drosophila*. *Cell.* **163**:1742–1755. (doi:10.1016/j.cell.2015.11.019).

[92] Aso Y, Sitaraman D, Ichinose T, Kaun KR, Vogt K, Belliart-Guérin G, Plaçais PY, Robie AA, Yamagata N, Schnaitmann C, Rowell WJ, Johnston RM, Ngo TT, Chen N, Korff W, Nitabach MN, Heberlein U, Preat T, Branson KM, Tanimoto H, Rubin GM. 2014. Mushroom body output neurons encode valence and guide memory-based action selection in Drosophila. *Elife* **3**:e04580. (doi: 10.7554/eLife.04580).

[93] Owald D, Waddell S. 2015. Olfactory learning skews mushroom body output pathways to steer behavioral choice in Drosophila. *Curr. Opin. Neurobiol.* **35**:178-84. (doi:10.1016/j.conb.2015.10.002).

[94] Eschbach C, Vogt K, Afonso B, Polizos N, Dancausse S, Evans A, Verbe A, Wang K, Berck M, Samuel A, Klein M, Cardona A, Zlatic M. 2025. A multisensory, bidirectional, valence encoder guides behavioral decisions. *Biorxiv.* (doi:10.1101/2025.09.26.678749).

[95] Rescorla RA, Wagner AR. 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF, editors. Classical conditioning II: current research and theory. New York: Appleton-Century-Crofts. p. 64–99.

[96] Van Hamme LJ, Wasserman EA. 1994. Cue competition in causality judgments: the role of nonpresentation of compound stimulus elements. *Learn. Motiv.* **25**:127–151. (doi:10.1006/lmot.1994.1008).

[97] Schultz W. 2015. Neuronal reward and decision signals: from theories to data. *Physiol. Rev.* **95**:853–951. (doi:10.1152/physrev.00023.2014).

[98] Bennett JEM, Philippides A, Nowotny T. 2021. Learning with reinforcement prediction errors in a model of the *Drosophila* mushroom body. *Nat. Commun*. **12**:2569. (doi:10.1038/s41467-021-22592-4).

[99] Jürgensen AM, Sakagiannis P, Schleyer M, Gerber B, Nawrot MP. 2023. Prediction error drives associative learning and conditioned behavior in a spiking model of *Drosophila* larva. *iScience*. **27**:108640. (doi:10.1016/j.isci.2023.108640).

[100] Stephen Maren S, K Luan Phan KL, Israel Liberzon I. 2013. The contextual brain: implications for fear conditioning, extinction and psychopathology. *Nat. Rev. Neurosci.* **14**:417-28. (doi:10.1038/nrn3492).
[101] Izquierdo I, Furini CRG, Myskiw JC. 2016. Fear Memory. *Physiol. Rev.* **96**:695-750. (doi: 10.1152/physrev.00018.2015).
[102] Roberts WA. 2019. The role of context in animal memory. *Learn. Behav.* **47**:117-130. (doi: 10.3758/s13420-019-00380-x).

Supplementary figures

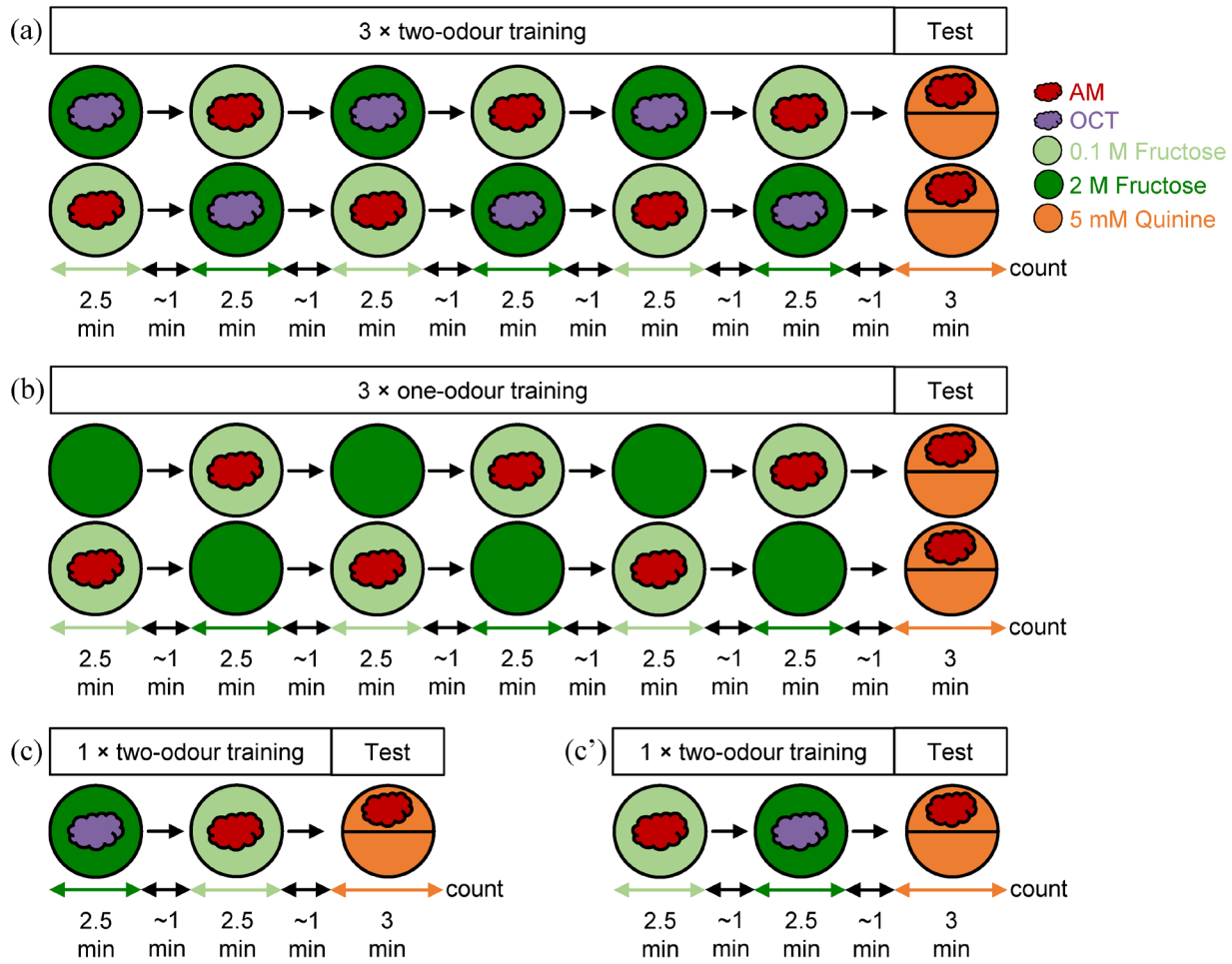

Figure S1. Training procedures in this study. (a) Larvae were trained for 3 cycles with two odours, AM and OCT and different concentrations of fructose. Depicted is group 6 in Table 1 as example, exactly as it was performed for Fig. 1. The larvae spent 2.5 minutes on each dish and were transferred manually with a brush between dishes, which took about 1 minute. After 3 complete cycles, the animals were transferred to a test dish covered with quinine and with only one odour on one side, and after 3 minutes the number of larvae on either side was counted. Each displayed training sequence was used in half of the repetitions of the experiment (samples). For Fig. S3a-c, the identities of AM and OCT were swapped; for Fig. S3d-f, the test dish contained pure agar instead of quinine; for Fig. 2, different quinine concentrations were used during training instead of fructose. (b) One-odour version of the experiment, as it was performed for Fig. 3a-c and S6. OCT was replaced by empty containers (not shown). For Fig. 3d-i, AM was replaced by OCT or BA in training and test. (c) One-trial version of the experiment as it was performed for Fig. 4a-c. (c') One-trial version of the experiment as it was performed for Fig. 4d-f.

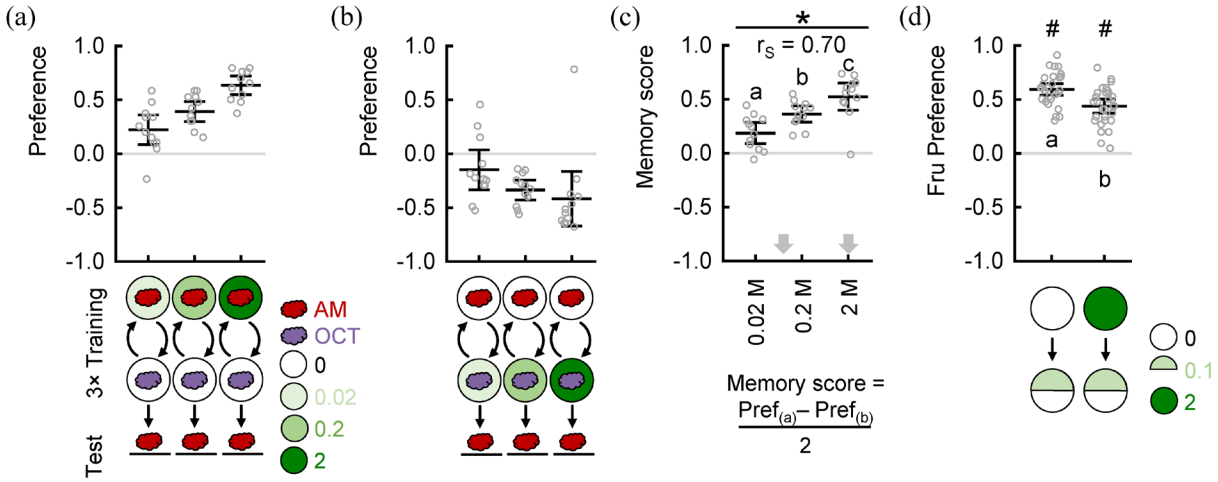

Figure S2. Reward strength increases with increasing fructose concentration. Larvae were trained with two odours, AM and OCT, and different concentrations of fructose (0.02, 0.2 and 2 M). (a) When AM had been paired with the fructose reward, the AM preference increased with increasing fructose concentration. (b) When OCT had been paired with the fructose reward, the AM preference decreased with increasing fructose concentration. (c) Memory scores were calculated as the difference in preference after both kinds of training. They significantly increased with increasing fructose concentrations. The arrows indicate the concentrations picked for the experiments in Figures 1, 3 and S3. (d) Larvae were exposed to either pure agar or agar mixed with 2 M fructose for 2.5 minutes. Subsequently, they were tested for their preference of a 0.1 M fructose substrate in a split Petri dish assay. In both cases, larvae preferred the fructose half. This demonstrates that the sensation and preference towards a low concentration of reward is intact after exposure to a high concentration of reward, as it is used in the learning experiments. Sample size is 12 each for (a-c), and 30 each for (d). Raw data are displayed with the mean ± 95 % confidence interval. In (c), * indicates a significant Spearman Rank test and r_S indicates Spearman's correlation coefficient. In (d), # indicates a significant Wilcoxon Signed Rank test against chance level (0). Different small letters above or below each group indicate significant pairwise difference from each other (Mann-Whitney U tests). For all raw data along with the detailed results of the statistical tests, see supplementary table S1.

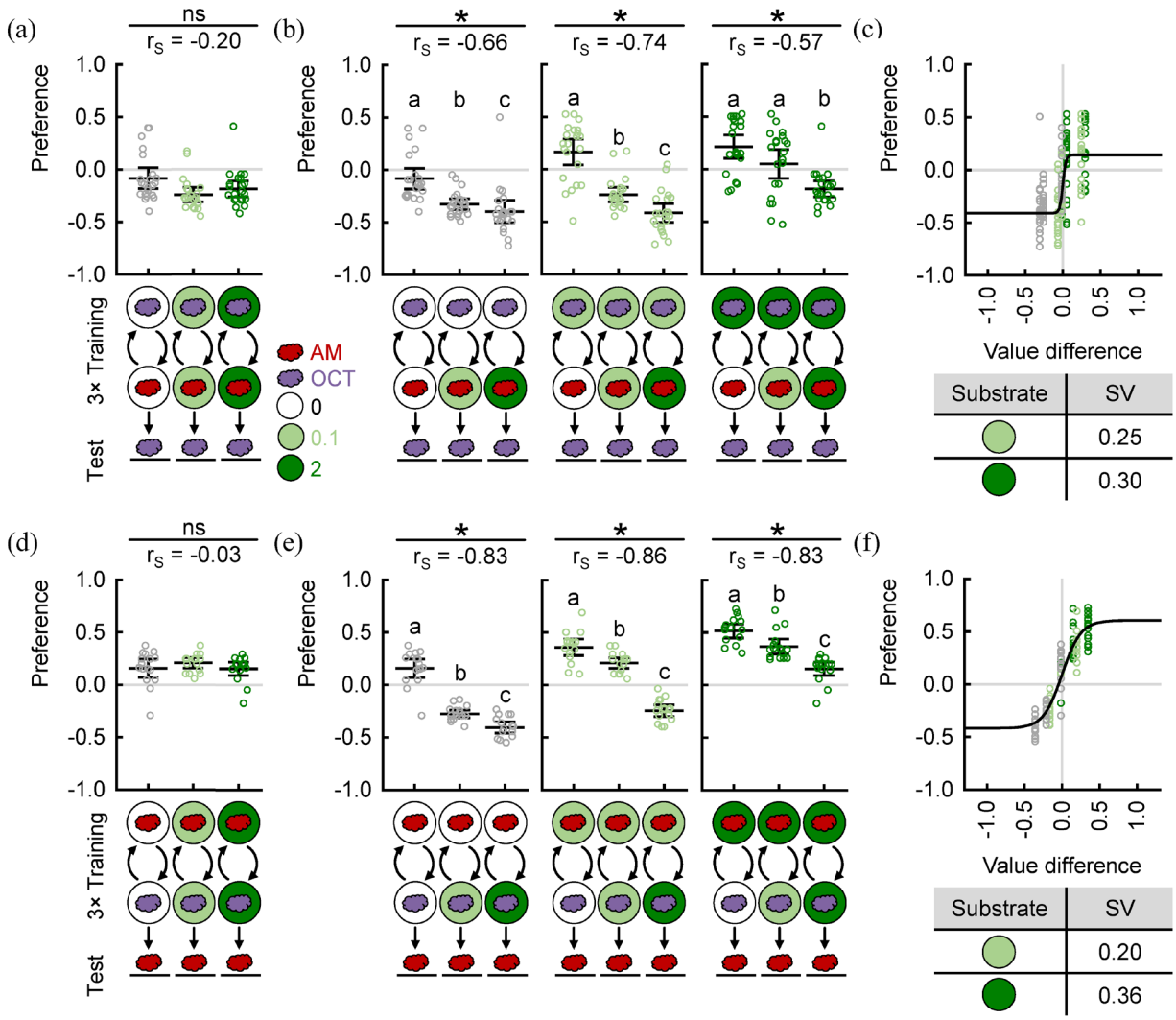

Figure S3. Larvae can learn the relative value of rewards independent of the tested odour or the test condition. (a-c) Larvae were trained with two odours, AM and OCT, and different concentrations of fructose. Afterwards, they were tested for their preference for OCT. (a) When the same fructose concentration was presented in both training dishes, the OCT preference was independent of the fructose concentration. (b) When the fructose concentration paired with OCT was kept constant, the OCT preference decreased with increasing fructose concentration paired with AM, and thus with increasing difference between the two rewards. (c) OCT preferences increased with increasing difference in reward value. The subjective values of each reward, estimated from the animals' behaviour, are displayed below the plot. (d-f) Same as (a-c), but the animals were tested for their AM preference and on a pure agar, tasteless Petri dish, in contrast to all other experiments that used a quinine test dish. The results were very similar to (a-c) and to Figure 1. Sample size is 22 each for (a-c), and 16 each for (d-f). Note that the groups displayed in (a) and (d) are also included in (b) and (e), respectively. Raw data are displayed with the mean ± 95 % confidence interval. * indicates a significant Spearman Rank test and r_S indicates Spearman's correlation coefficient. Different small letters above each group indicate significant pairwise difference from each other (Mann-Whitney U tests). For all raw data along with the detailed results of the statistical tests, see supplementary table S1.

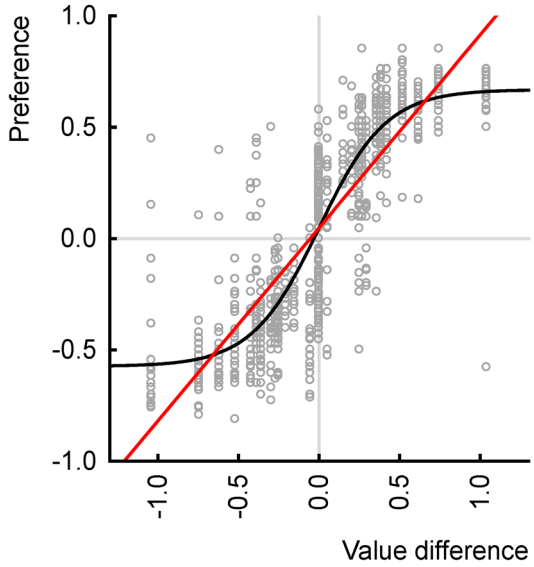

Figure S4. The relationship between the value difference and the odour preference is logistic. This graph displays all data of this study using three training trials with two odours (Fig. 1, 2, S3). To compare a generalized logistic function with a simple linear function, we calculated the Akaike information criterion (AIC) that is regularly used to assess the goodness of a fitting. The AIC for the generalized logistic model was lower than the AIC for the linear function (-390.1 versus -133.9), indicating that the generalized logistic model fitted the data better than the linear one.

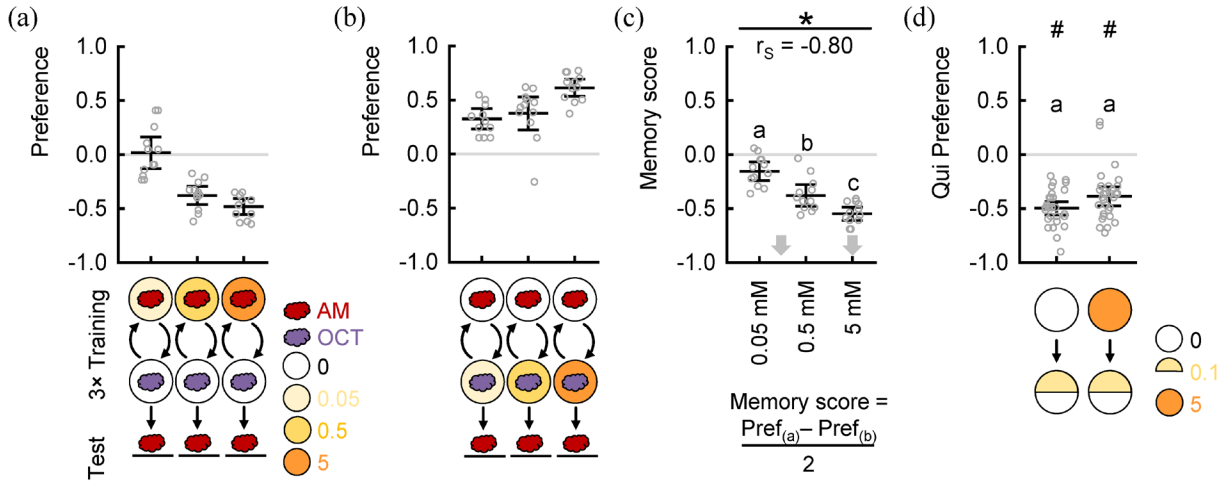

Figure S5. Punishment strength increases with increasing quinine concentration. Larvae were trained with two odours, AM and OCT, and different concentrations of quinine (0.05, 0.5 and 5 mM). (a) When AM had been paired with the quinine punishment, the AM preference decreased with increasing quinine concentration. (b) When OCT had been paired with the quinine punishment, the AM preference increased with increasing quinine concentration. (c) Memory scores were calculated as the difference in preference after both kinds of training. They got significantly more negative with increasing quinine concentrations. The arrows indicate the concentrations picked for the experiments in Figure 2. (d) Larvae were exposed to either pure agar or agar mixed with 5 mM quinine for 2.5 minutes. Subsequently, they were tested for their preference of a 0.1 mM quinine substrate in a split Petri dish assay. In both cases, larvae avoided the quinine half. This demonstrates that the sensation and repulsion towards a low concentration of punishment is intact after exposure to a high concentration of punishment, as it is used in the learning experiments. Sample size is 12 each for (a-c), and 30 each for (d). Raw data are displayed with the mean ± 95 % confidence interval. In (c), * indicates a significant Spearman Rank test and r_S indicates Spearman's correlation coefficient. In (d), # indicates a significant Wilcoxon Signed Rank test against chance level (0). Different small letters above each group indicate significant pairwise difference from each other (Mann-Whitney U tests). For all raw data along with the detailed results of the statistical tests, see supplementary table S1.

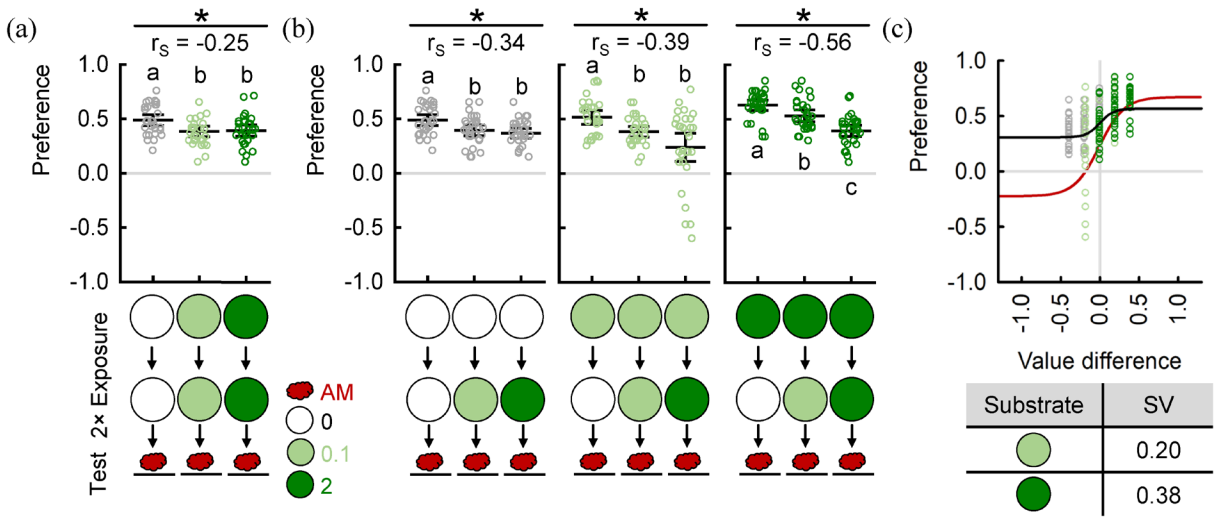

Figure S6. Sugar exposure affects odour preference in a sequence-dependent manner. Larvae were sequentially exposed to two dishes containing various concentrations of fructose. Afterwards, they were tested for their preference for AM. (a) When the same fructose concentration was presented in both exposure dishes, the AM preference was decreasing after fructose exposure relative to the baseline group. (b) When the fructose concentration in the first dish was kept constant, the AM preference decreased with increasing fructose concentration in the second dish. (c) AM preferences increased with increasing difference in reward value, which was calculated in this case as the reward value of the first minus the second exposure dish. The subjective values of each reward were estimated from the animals' behaviour in the experiment displayed in Fig. 4d-f which had been performed in parallel. In addition to the generalized logistic function fitted to the displayed data (black), the corresponding fit of Fig. 4f is shown (red) for easy comparison. Sample size is 30 each. Note that the groups displayed in (a) are also included in (b). Raw data are displayed with the mean ± 95 % confidence interval. * indicates a significant Spearman Rank test and r_S indicates Spearman's correlation coefficient. Different small letters above or below each group indicate significant pairwise difference from each other (Mann-Whitney U tests). For all raw data along with the detailed results of the statistical tests, see supplementary table S1.