# 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 use 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 realized even by simple insect brains. Given the great potential of *D. melanogaster* as a model organism for in-depth neurobiological analyses, our study opens up the opportunity to reveal the mechanism underlying relative value learning in unprecedented detail.

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 other options [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 experience 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 do 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, 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]. Studies in other mammals and in birds suggest that these principles are shared across vertebrates [16-22].

Compared to vertebrates, invertebrates like insects have very differently structured brains that are much smaller and often regarded as too simple to allow for complex tasks. Although associative learning has been studied in insects like bees, crickets and flies since decades [23-25], the question whether insects evaluate rewards or punishments relatively has rarely been asked. Studies in locusts, bees, ants and fruit flies have demonstrated that insects do discriminate values of different rewards or punishments and can learn accordingly [26-33]. However, we are not aware of any demonstration that insects actually learn about, that is, “save” the relative value of a reward or punishment rather than the absolute value. Do insects learn about relative values of rewards and punishments like vertebrates, or do they use absolute values?

Here, we use the larval stage of the fruit fly *Drosophila melanogaster*. In the past decades, the *D. melanogaster* larva has established itself as extremely useful model organism for neurobehavioural research that often pioneers and inspires research in animals with more complex brains [34-38]. It combines a simple brain with only about 10,000 neurons, almost unlimited possibilities for transgenic manipulations of individual cells, and a surprisingly complex behavioural repertoire given its simplicity. This includes appetitive and aversive classical as well as operant conditioning, long-term-memories, decision-making and adaptive outcome expectations [26, 39-48].

We designed a behavioural paradigm to distinguish absolute and relative value learning in classical conditioning. Based on established odour-taste associative experiments [39, 49-51], we systematically ask whether these simple animals learn about the relative values of rewards as well as punishments and whether this learning depends on the odours presented during the memory retrieval test, the number of training trials and the sequence of training trials. Our findings will provide an opportunity to investigate the detailed neuronal mechanism of relative value learning in a 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 [40], 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++), 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). 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.

$$\text{Preference} = (\#AM - \#EM) / (\#AM + \#middle + \#EM)$$

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

[50]. 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:

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

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:

$$SV_{\text{neutral}} = 0$$

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

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

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 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+) or AM was presented first (AM+ → OCT++).

*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.

| | | 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.

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 [40]. 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.

*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:

$$\text{Fru Preference} = (\#\text{Fru} - \#\text{Pure}) / (\#\text{Fru} + \#\text{middle} + \#\text{Pure})$$

$$\text{Qui Preference} = (\#\text{Qui} - \#\text{Pure}) / (\#\text{Qui} + \#\text{middle} + \#\text{Pure}) .$$

*Statistical analyses*

Data are displayed as raw data points, overlaid with the mean as middle line and the 95 % confidence interval as whiskers. 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 % [52].

In experiments measuring taste preference, we determined whether the preferences were significantly different from zero using the Wilcoxon Signed-Rank test (WSR).

To determine whether gradual increases of reward or punishment strength consistently affects odour preferences and memory scores, Spearman Rank Correlations (SRC) were performed. 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). Graphs, WSR and SRC tests were realized in Statistica 11 (StatSoft, Tulsa, USA).

As the preference scores in our experiments are technically limited between -1 and 1, their relationship with any of our experimental parameters 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 [53]. The function followed the form:

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

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 [54], and are reported in Supplementary table S1 along with the raw data. To compare the generalized logistic function with a simpler linear function, we compared the Akaike information criterion (AIC) that balances the likelihood of a fitted model with its complexity. The AIC is regularly used to compare statistical models for the same data set, with the model with the lowest AIC being considered the best fit [55].

## Results

*Larvae learn about the relative value of rewards*

We performed an associative learning experiment that was modified from the most commonly used learning paradigm in *D. melanogaster* [50, 56]. We trained larvae by alternatingly pairing *n*-amylacetate (AM) with a given concentration of fructose reward and 1-octanol (OCT) without reward, and subsequently tested them for their preference of AM alone (Fig. 1a). The presence of only AM during the test, different from the standard procedure, allowed us to assess what the animals had learned specifically about this odour. Two fructose concentrations were chosen based on initial experiments (Fig. S1a-c). We ensured that the larvae can still detect and process the lower of the two fructose concentrations immediately after being exposed to the higher concentration (Fig. S1d).

Larvae have an innate preference to odours that can be modulated by non-associative processes [49, 51]. We therefore included a baseline group with the same handling and odour exposure as all other groups, but without any rewards and therefore arguably without associative learning (Fig. 1a, left-most group). Compared to this baseline, we found that the animals' preference for AM increased with increasing reward concentration, suggesting increased associative learning (Fig. 1a, see also Fig. S1a-c).

Importantly, in this experiment the larvae could learn to associate AM either with an absolute value of the paired reward (how good it was), or with its relative value (how much better it was

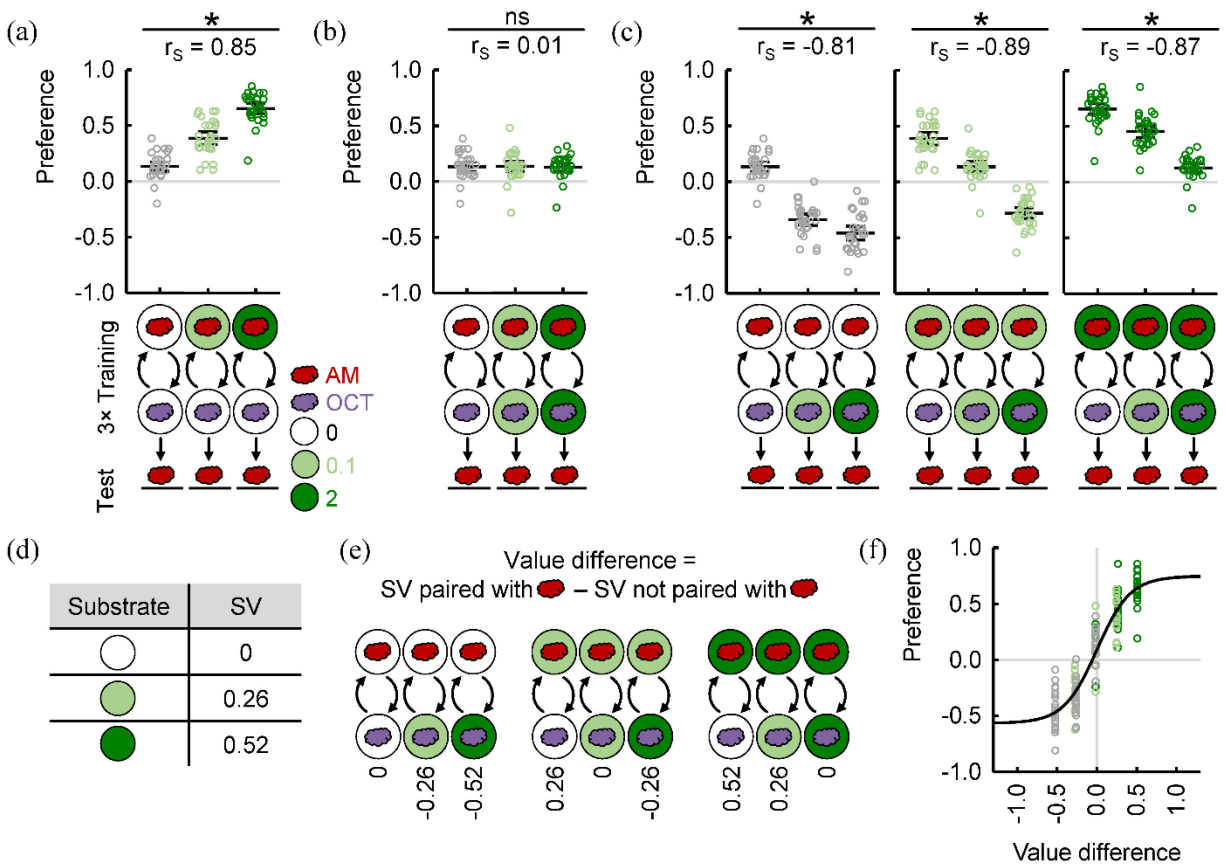


**Figure 1**. Larvae learn about the relative value of rewards. (a) Larvae were trained with AM paired with 0, 0.1 or 2 M fructose reward and OCT paired with no reward. (b) Larvae were trained with the same fructose concentration in both training dishes. (c) Larvae were trained with the fructose concentration 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. For all raw data along with the detailed results of the statistical tests, see supplementary table S1.

195 than the alternative). As the alternative during training was plain, neutral agarose, both hypotheses
196 predict the same outcome and the current experiment does not allow to differentiate between them.
197 Therefore, we trained the larvae with AM and OCT being paired with the same fructose
198 concentration (Fig. 1b). If the animals learned about the absolute reward value, the substrate that
199 was paired with OCT should not matter and the animals should show the same level of associative
200 learning as in Fig. 1a. If the animals learned about the relative reward value, however, there should
201 be no sign of associative learning in this experiment, because both alternating substrates were the

202 same and therefore the relative value of the reward paired with AM was zero (neither better nor
203 worse than the alternative). Indeed, we found that the animals all displayed the same baseline AM
204 preference, irrespective of the reward strength (Fig. 1b). This suggests that they learned about the
205 relative and not the absolute reward value.

206 Next, we modified the experiment such that AM was paired with no reward, and thus the
207 absolute reward value was zero, but the alternative featured different reward concentrations (Fig.
208 1c, left). We found that the AM preference was decreasing with increasing reward strength that
209 had been paired with OCT. Again, this fits to the hypothesis of relative value learning: as the
210 substrate paired with AM had been less good than the alternative, its relative value was negative.
211 Finally, we repeated the same experiment, but paired AM either with the weak reward (Fig. 1c,
212 middle) or the strong reward (Fig. 1c, right). In each case, the AM preference declined along with
213 the reward strength that had been paired with OCT, and in each case the result indicated that the
214 animals' behaviour was determined by the relative reward value associated with AM. Strikingly,
215 this even caused the larvae to avoid AM after it had been paired with a reward, but the alternative
216 substrate had been even better (Fig. 1c, 6$^{th}$ group from the left). Thus, we concluded that *D.*
217 *melanogaster* larvae learn about the relative value of a reward, not the absolute value. We obtained
218 an analogous result when we repeated the complete experiment but tested for the OCT preference
219 instead of AM (Fig. S2a-c).

220 In this study we used both fructose reward and quinine punishment (see below). Quinine
221 memories have been repeatedly shown to be only retrieved when the animals are tested in presence
222 of quinine [26, 40, 57-59]. Therefore, and in order to ensure comparability throughout the study,
223 we used 5 mM quinine test Petri dishes for all experiments, including experiments that only used
224 reward during training. Previous work has demonstrated that the retrieval of fructose reward
225 memories is not influenced by the presence of quinine [40]. To further control whether the test
226 condition does not influence our the results, we repeated the complete experiment once more, but
227 with the test performed on a pure, tasteless Petri dish, and obtained very similar results irrespective
228 of the test condition (Fig. S2d-f).

229 To illustrate how the difference in value between the two substrates determines the preference,
230 we deduced the subjective values of the two rewards directly from the animals' behaviour during
231 the experiment in Fig. 1a, as the average preference of a group that received AM-reward training
232 minus the average of the baseline preference group (Fig. 1d; for more details, see Materials &

233 Methods). Then, we determined the value difference in each of the experimental groups of Fig. 1c
234 by subtracting the subjective value of the substrate paired with OCT from the value of the substrate
235 paired with AM (Fig. 1e). Finally, we plotted the AM preferences of all groups over the value
236 differences and used a generalized logistic function to describe the relationship between them (Fig.
237 1f, Fig. S2c,f). Across the data of this study, a generalized logistic function fits the data better than
238 a linear function (Fig. S3).
239 In summary, we found that larvae learned about the relative value of rewards, and that the value
240 difference between the options during training determined the odour preference in a logistic
241 manner.
242
243 *Larvae learn about the relative value of punishments*
244 Next, we asked whether the larvae learn about the absolute or relative value of punishments. To
245 answer this, we performed a similar experiment as shown in Figure 1, but used two different
246 concentrations of quinine as punishment, based on an initial experiment (Fig. S4).
247 Similar to the previous findings, the animals all displayed the same baseline AM preference
248 after they had been trained with the same quinine concentration in both training dishes, irrespective
249 of the punishment strength (Fig. 2a). When AM was paired with no punishment, and thus the
250 absolute punishment value was zero, but the alternative featured different punishment
251 concentrations (Fig. 2b, left), we found that the AM preference was increasing with increasing
252 punishment strength that had been paired with OCT. Finally, when AM had been paired either with
253 the weak punishment (Fig. 2b, middle) or the strong punishment (Fig. 2b, right), in each case the
254 AM preference increased along with the punishment strength that had been paired with OCT. Thus,
255 we concluded that larvae learn about the relative value of a punishment, not the absolute value.
256
257 *Larvae learn about the relative values of combined reward and punishment*
258 After finding that the larvae learned about the relative value of rewards and punishments in
259 separate experiments, we wondered what would happen if we combined reward and punishment
260 in the same experimental setup. When both odours had been paired with the same reward or
261 punishment, we saw a slight but significant difference in AM preference, mostly due to a decreased
262 AM preference after training with quinine (Fig. 2d). We note that this group is an exact repetition
263 of the one in Fig. 2a (right-most group) in which we saw no reduced preference. When keeping

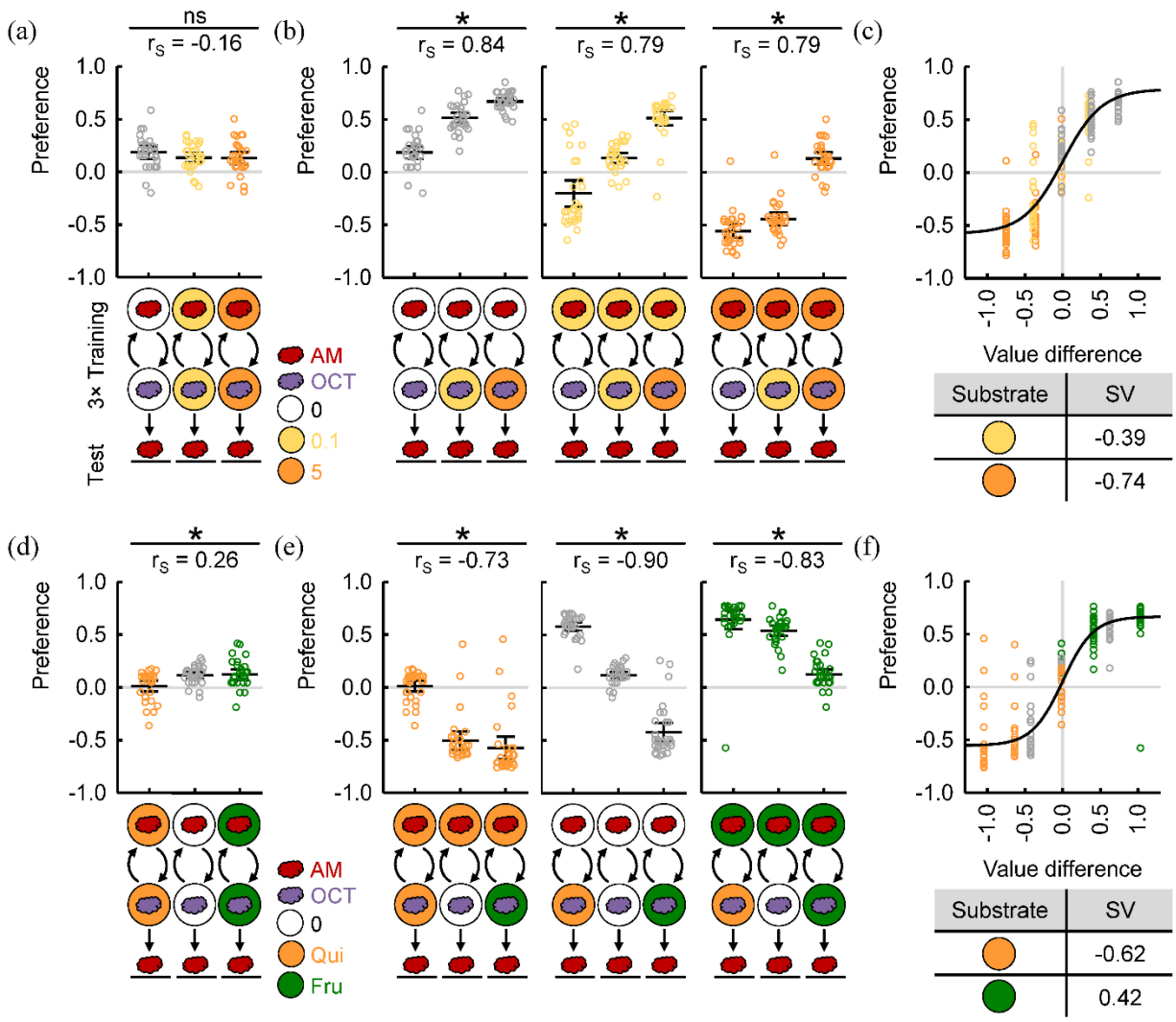


**Figure 2**. Larvae learn about the relative value of punishments. (a-c) Larvae were trained with two odours, AM and OCT, and different concentrations of quinine (0, 0.1 and 5 mM). (a) The same quinine concentration was presented in both training dishes. (b) The quinine concentration paired with AM 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.

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 to stronger negative or positive preferences than training with only the reward or

268 only the punishment (Fig. 2e, 3rd group from the left and from the right). Thus, we again concluded
269 that larvae learn about the relative values of reward and punishment, not the absolute values. In
270 this experiment, as the value differences between the substrates were particularly large, the logistic
271 relationship (as compared to a simple linear one) between the value difference and the odour
272 preference was particularly striking (Fig. 2f).
273
274 *Larvae learn relatively independent of the training sequence*
275 Naturally, we wondered how the larvae compared the values of the different training
276 concentrations in order to learn about the relative value of each substrate. We first hypothesized
277 that the animals monitored the relative change of reward strength over time and associated this
278 change with the stimuli currently presented. Thus, an increase in reward strength at the moment
279 AM was presented established an appetitive memory for AM, whereas a decrease established an
280 aversive memory. Similar learning processes have been observed before in many species,
281 including in *D. melanogaster* [60-62]. Crucially, this process can only work if AM is presented
282 after the alternative substrate. If AM is presented first, the larvae have not experienced the
283 alternative substrate yet and thus cannot associate any change in reward strength with AM.
284 To test this hypothesis, we switched to a one-trial training paradigm [63]. The training was
285 performed as described for Fig. 1, but the larvae were exposed to each odour-substrate combination
286 only once before they were tested for their AM preference. This allowed us a precise control over
287 the training sequence without the complication of repeated training cycles. We first trained the
288 animals such that OCT was always presented first and AM second (Fig. 3a-c). The results were
289 very similar to the three-trial experiments and clearly indicated relative value learning. Then, we
290 trained the animals with the other sequence, presenting AM first and OCT second (Fig. 3d-f). When
291 both odours had been paired with the same reward strength, we found a slight but significant
292 increase of the AM preference with increasing reward strength (Fig. 3d). Surprisingly, when we
293 kept the substrate paired with AM constant, the AM preference decreased with increasing reward
294 strength paired with OCT, indicating relative value learning (Fig. 3e). Overall, the results were
295 very similar between both training sequences. This contradicts our initial hypothesis and suggests
296 that the memory associated with AM was retrospectively modified by the reward strength that was
297 presented after the AM presentation had ended. Thus, larvae can learn about the relative reward
298 value irrespective of the training sequence.

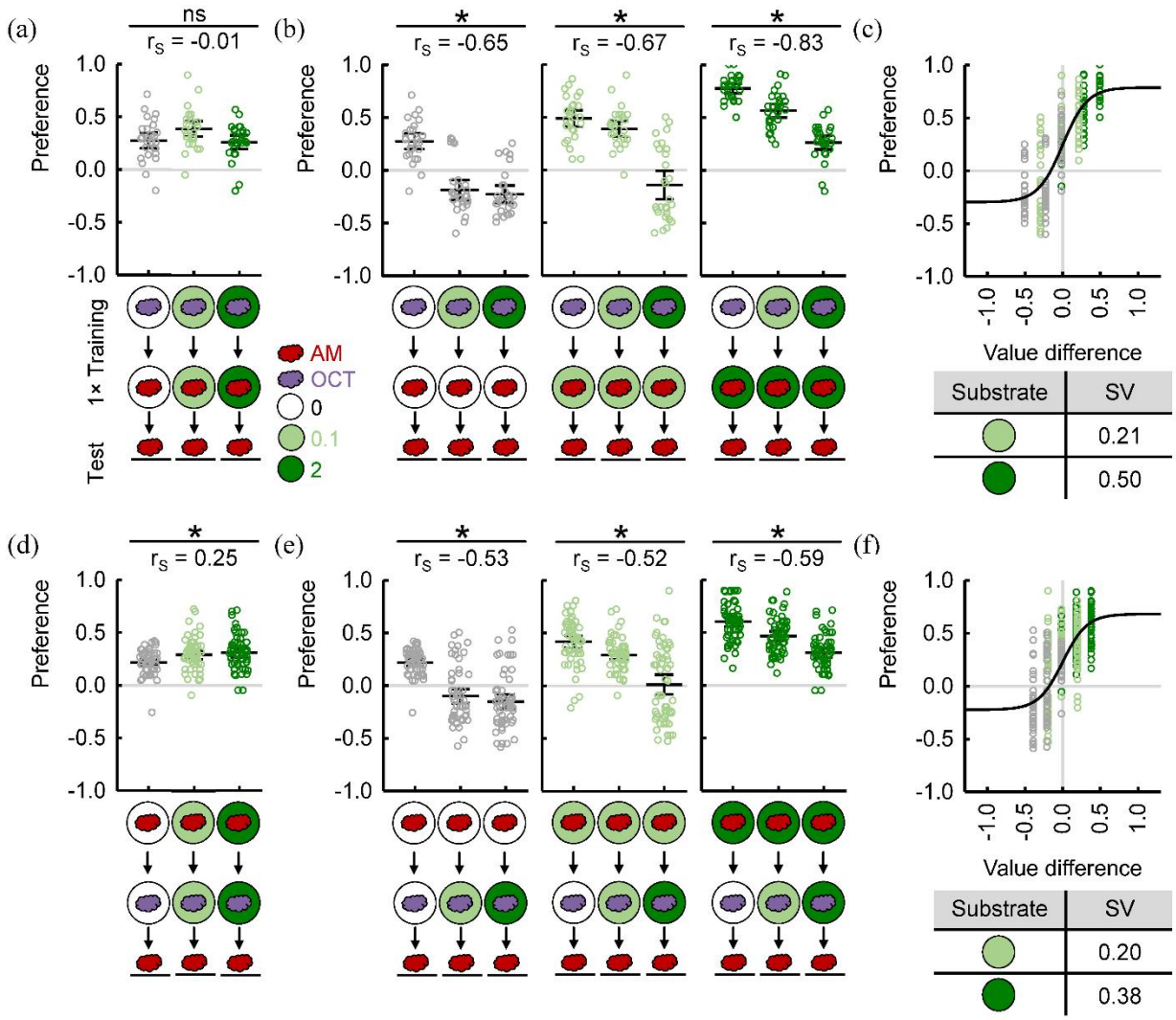


**Figure 3**. Larvae can learn the relative value of rewards independent of training sequence. Larvae were trained with two odours, AM and OCT, and different concentrations of fructose. Only a single training trial was performed, with (a-c) OCT or (d-f) AM being presented first. (a, d) The same fructose 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. We found that larvae employ relative value learning both for rewards (Fig. 1) and punishments (Fig. 2), irrespective of which odour was tested (Fig. S2) as well as of the number and sequence of training trials (Fig. 3). Throughout these experiments, we found no clear evidence that the absolute value was learned at all.

It is 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]. We found a similar kind of suboptimal decision in our experiments: 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 (Fig. 1c, S2b,e, 6$^{th}$ group from the left). This prevented the larvae to find an otherwise positively rated sugar source. Thus, it seems that the way how brains evaluate rewards or punishments during learning is shared across the animal kingdom from insect larvae to humans.

*Relative value learning is consistent with previous studies*

To our knowledge, this is the first demonstration of relative value learning in *D. melanogaster*. Previous studies in both larvae and adult flies had used different punishment strengths during training, each paired with a given odour, and then had tested the animals' choice between these odours [26, 29, 33]. These studies demonstrated that the animals can learn how bad a given punishment is, but they could not distinguish between relative and absolute value learning, as it remains possible that the animals learned the absolute value paired with each odour and then compared during the choice test which of the odours predicted the better outcome. The critical difference to our design is that we tested only a single odour and therefore could directly assess what the animals had learned about it.

A number of studies used an experimental paradigm similar to ours with two odours during the training but only one of them during the test [51, 64-66]. In addition, a one-odour training version

330 is widely used in larvae [36, 40, 48-49, 59, 67-72]. All of these studies used only a single reward
331 or punishment paired with one odour, while the other training situation was neutral and therefore
332 had a value of zero. Nevertheless, their results are consistent with the hypothesis of relative value
333 learning.

334 For example, one-odour experiments with rewards always consisted of two types of training
335 regimen: either the odour was paired with the reward while the other training situation featured no
336 odour on a neutral substrate (so-called paired training), or the odour was paired with the neutral
337 substrate while the reward was presented without an odour (so-called unpaired training). For paired
338 training, both absolute and relative value learning would lead to the same outcome, an increased
339 odour preference compared to the baseline. For unpaired training, in contrast, absolute value
340 learning would lead to a baseline level of odour preference because the odour had not been paired
341 with any reward, whereas relative value learning would lead to a preference lower than the baseline
342 because the odour had been paired with the less rewarding of the two substrates. The latter is
343 precisely what had been found in several previous studies [48-49, 51, 59]. Analogous results have
344 been observed when using punishments instead of rewards [51, 59] as well when using a training
345 with two odours (Fig. 1c, 2b) [51]. Using a different experimental design, similar results have been
346 found in adults, too [65].

347 These observations gave rise to the proposal of two separate learning processes after paired and
348 unpaired training, leading to memories of opposite valences [49, 51, 59, 65, 73]. However, relative
349 value learning provides a unified framework to explain the results of both kinds of training: the
350 memory's valence depends on whether the trained odour had been paired with the better or worse
351 of the training situations.

352

353 *Practical implications*

354 Our results have significant implications for interpreting the results of other studies on associative
355 learning in *D. melanogaster*. Our experiments followed the standard procedures that have been
356 used in most previous works in the larva [50, 56]. It is therefore very likely that in most previously
357 published experiments the larvae were learning about the relative values of rewards and
358 punishments, too. Furthermore, in adult flies the very same experimental logic is used in the
359 overwhelming majority of studies, with two odours presented alternately, one of them paired with
360 a reward or punishment [74-77]. Thus, it is quite plausible that relative value learning occurred in

many of those studies, too. This implies that the current theoretical, computational and neuronal working models of associative learning in *D. melanogaster* are lacking this important aspect and should be carefully revised.

In the following, we will briefly discuss how relative value learning could be plausibly integrated into the current theoretical and neuronal working models.

*A potential learning theory for relative value learning*

For relative value learning, the animals need to compare the reward (or punishment) strengths in presence and absence of an odour, and associate their difference with the odour. Interestingly, one of the most influential and widely adapted theories for associative learning may provide a simple way to accomplish this. The Rescorla-Wagner model [78] states that associative learning ($\Delta V$) is driven by a prediction error, that is, the difference between the associative strength supported by a reward currently present ($\lambda$) minus the already existing associative strength from previous memories for all stimuli that are currently present ($\sum V$). In simplified terms, the prediction error can be seen as the difference between the current reward and previously experienced rewards.

The prediction error can serve as a tool to compare between reward strengths in presence and absence of the odour in our case if we assume a context memory. Given how significantly different the experimental context (potential background odour, light, substrate texture) is from the usual living conditions of the larvae, it is plausible that they associate some contextual cues with the rewards during the training. Consider the following training sequence: OCT++ → AM+ (Fig. 3b, 6$^{th}$ group from the left). In the first training phase, the animals form an association between the strong reward and the experimental context. In the second phase, the presence of the same context predicts a strong reward, yet only a weak reward is present, leading to a negative prediction error ($\lambda < \sum V$). This negative prediction error is associated with AM, leading to the observed AM avoidance (Fig. 3b, 6$^{th}$ group from the left).

Now consider the opposite training sequence, AM+ → OCT++ (Fig. 3e, 6$^{th}$ from the left): in the first training phase, a positive prediction error is associated with AM and the context. In the second phase, more reward is present than predicted by the context, resulting in another positive prediction error. How can a positive prediction error lead to an aversive memory for AM as we observed it? A simple extension of the Rescorla-Wagner model could provide an explanation [79]. This model had been developed to account for a variety of observations in mammals which are

summarized under the term retrospective revaluation [80]. Applied to our experiment, the model assumes that in the first training phase the odour and the context are not only associated with the reward but also with each other. In the second training phase, the odour is predicted by the context but missing. This inverts the sign of the formed association such that a positive prediction error causes a negative association with the absent odour. In plain language, the animals learn that the context alone predicts more reward than the context and the odour together, and therefore retrospectively make the memory regarding the odour more negative.

Thus, prediction-error theories provide a possible explanation for relative value learning. How would a simple neuronal circuit realizing these theories look like?

*A potential neuronal circuit for relative value learning*

A hypothetical neuronal circuit that explains our results obviously requires neurons coding for the reward, the context and the odour, and at least three components (Fig. 4): (1) a mechanism to associate the context and odour pathways with each other which in the simplest case could be accomplished by reciprocal synapses; (2) a mechanism to associate both the context and odour pathways with the reward signal; and (3) a neuronal pathway that gathers the output of those associations, representing $\sum V$, and feed it back to the reward pathway in order to allow the calculation of a prediction error.

Considering the training sequence OCT++ → AM+, the coincidence of the context with the strong reward in the first training phase establishes a strong associative memory that is retrieved in the second phase. Because the reward in the second phase is smaller than the predicted reward, the prediction error is negative, resulting in a below-threshold activity of the reward pathway and an aversive memory associated with AM.

Considering the opposite training sequence (AM+ → OCT++), the coincidence of the context, AM and the weak reward in the first training phase establishes reciprocal associations between the context and AM as well as associations of both of them with the reward. In the second phase, the presence of a stronger reward than predicted by the context leads to a positive prediction error and an above-threshold activity in the reward pathway. Critically, because AM is absent but predicted by the context-AM associations, the positive prediction error results in a negative change of associative strength between AM and the reward.

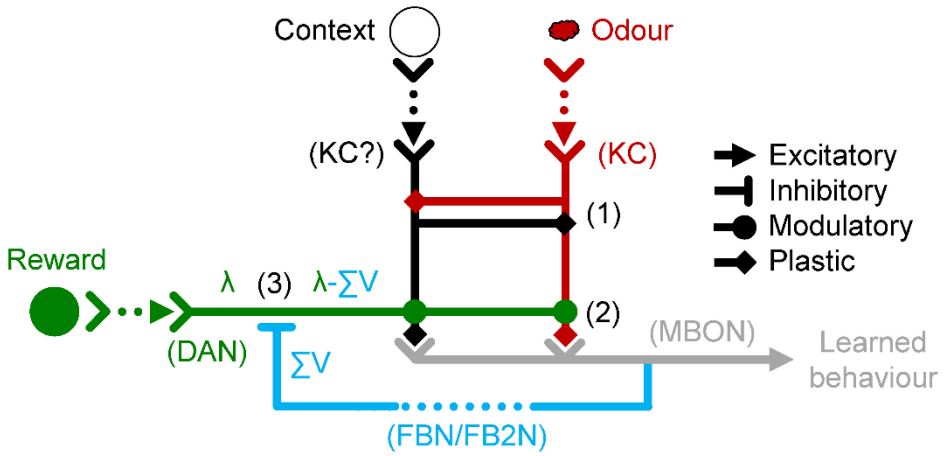


**Figure 4**. Simplified potential circuitry underlying relative value learning. Neuronal pathways coding for stimuli like context or odour are interconnected by reciprocal plastic synapses (1). Coincidence of activity of a stimulus pathway with the reward pathway modulates plastic synapses to a shared output pathway (2). Critically, the sign of the synaptic modulation is opposite when a stimulus is absent but predicted by another stimulus via the reciprocal plastic synapses. The output pathway collects memory information from all active stimulus pathways (∑V) and feeds it back to the reward pathway where it gets subtracted from the current reward signal (λ) (3). Candidate neurons of the mushroom body circuitry for each hypothetical neuronal pathway are indicated in brackets. DAN: dopaminergic neuron; KC: Kenyon cell; MBON: mushroom body output neuron, FBN: feedback neuron; FB2N: two-step feedback neuron. For simplicity, only a single mushroom body compartment is shown; additional synaptic connections, in particular between-compartment connections, are omitted. Based on [36, 68, 70].

Interestingly, one of the major memory centres in insects, the mushroom body, can be mapped very well to this hypothetical circuit (for reviews in *D. melanogaster*, see [25, 81]). In brief, odours and other stimuli are coded by specific combinations of Kenyon cells (KCs), the intrinsic neurons of the mushroom body, which are strongly interconnected (Fig. 4, "1") [36, 82]. To explain our results, we need to assume that the contextual cues are coded among the KCs and that the reciprocal KC-KC connections are plastic, which remains to be tested.

The KCs are intersected by (mostly) dopaminergic modulatory input neurons (DANs) and mushroom body output neurons (MBONs). The DANs receive signals from rewards (or punishments) and synapse to the axonal terminals of all KCs. The coincidence of stimulus-activated KCs and DAN signals modulates the KC-MBON synapses (Fig. 4, "2"). The synaptic weights of the KC-MBON synapses thus contain the associative strengths of all stimuli coded by the KCs.

The MBONs gather input from all KCs (and thus code for the sum of the associative strengths of all present stimuli, ∑V) and route back via one and two step feedback relays to the DANs [70]. This allows the DANs to integrate current sensory reward signals (λ) with the mushroom body's reward prediction. In other words, this circuit enables the DANs to calculate a prediction error, analogous to what has been found in mammals [83] (Fig. 4, "3"). Current computational models of *D. melanogaster* associative learning implement the prediction error in such way and use it to explain higher-order phenomena that require prediction-dependent teaching signals [70, 72-63, 84-86].

To explain our results, we need to assume that the same dopaminergic signal can result in opposite modulations of the KC-MBON synapses when a KC is active, or when it is inactive but its activity is predicted by the reciprocal KC-KC synapses. Although such a mechanism is speculative at this point, we note that it has been shown that the same dopaminergic signal can have opposite effects on the KC-internal signalling pathways dependent on the activity state of the KCs [87]. In addition, the KC-MBON synapses can be both potentiated and depressed depending on training contingencies [88].

In summary, the mushroom body circuit provides a good candidate substrate to accomplish relative value learning in insects, but critical assumptions need yet to be tested. In any case, the *D. melanogaster* larva, with its small brain and excellent tools for transgenic manipulations, is perfectly suited to reveal the detailed mechanism and to deepen our understanding of how brains integrate the relative values of options into associative learning.

## 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] Belke TW. 1992. Stimulus preference and the transitivity of preference. *Anim. Learn. Behav*. **20**:401–406. (doi:10.3758/BF03197963).

[17] 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).

[18] Waite TA. 2001. Background context and decision making in hoarding gray jays. *Behav. Ecol*. **12**:318–324. (doi:10.1093/beheco/12.3.318).

[19] Tobler PN, Fiorillo CD, Schultz W. 2005. Adaptive coding of reward value by dopamine neurons. *Science*. **307**:1642–1645. (doi:10.1126/science.1105370).

[20] 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).
[21] 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).
[22] 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).
[23] Giurfa M. 2015. Learning and cognition in insects. *Wiley Interdiscip. Rev. Cogn. Sci.* **6**:383-395. (doi: 10.1002/wcs.1348).
[24] 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).
[25] 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).
[26] 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).
[27] 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).
[28] Gil M, De Marco RJ, Menzel R. 2007. Learning reward expectations in honeybees. *Learn. Mem.* **14**:491–496. (doi:10.1101/lm.618907).
[29] 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).
[30] 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).
[31] 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).
[32] 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).
[33] 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).
[34] Sokolowski MB. 2001. *Drosophila*: genetics meets behaviour. *Nat. Rev. Genet.* **2**:879–890. (doi:10.1038/35098592).
[35] 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).
[36] 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).

[37] 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).

[38] 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).

[39] Scherer S, Stocker RF, Gerber B. 2003. Olfactory learning in individually assayed *Drosophila* larvae. *Learn. Mem*. **10**:217–225. (doi:10.1101/lm.57903).

[40] 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).

[41] 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).

[42] 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).

[43] 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).

[44] 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).

[45] 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).

[46] 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).

[47] 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).

[48] 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).

[49] 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).

[50] 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).

[51] 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).

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

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

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

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

[56] 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).

[57] 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).

[58] 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).

[59] 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).

[60] Gerber B, Yarali A, Diegelmann S, Wotjak CT, Pauli P, Fendt M. 2014. Pain-relief learning in flies, rats, and man: basic research and applied perspectives. *Learn. Mem.* **21**:232–252. (doi:10.1101/lm.032995.113).

[61] Weiglein A, Thoener J, Feldbruegge I, Warzog L, Mancini N, Schleyer M, 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. (doi:10.1002/cne.25037).

[62] 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).

[63] 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).

[64] 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).

[65] 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).

[66] 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).

[67] 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).

[68] 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).

[69] 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).

[70] 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).

[71] 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).

[72] 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).

[73] 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).

[74] Tempel BL, Bonini N, Dawson DR, Quinn WG. 1983. Reward learning in normal and mutant *Drosophila*. *Proc. Natl Acad. Sci. USA*. **80**:1482–1486. (doi:10.1073/pnas.80.5.1482).

[75] Tully T, Quinn WG. 1985. Classical conditioning and retention in normal and mutant *Drosophila melanogaster*. *J. Comp. Physiol. A.* **157**:263–277. (doi:10.1007/BF01350033).

[76] Krashes MJ, Waddell S. 2011. *Drosophila* appetitive olfactory conditioning. *Cold Spring Harb. Protoc*. **2011**:pdb.prot5609. (doi:10.1101/pdb.prot5609).

[77] Krashes MJ, Waddell S. 2011. *Drosophila* aversive olfactory conditioning. *Cold Spring Harb. Protoc*. **2011**:pdb.prot5608. (doi:10.1101/pdb.prot5608).

[78] 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.

[79] 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).

[80] Miller RR, Witnauer JE. 2016. Retrospective revaluation: the phenomenon and its theoretical implications. *Behav. Process*. **123**:15–25. (doi:10.1016/j.beproc.2015.09.001).

[81] 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).

[82] Schürmann FW. 2016. Fine structure of synaptic sites and circuits in mushroom bodies of insect brains. *Arthropod Struct. Dev*. **45**:399–421. (doi:10.1016/j.asd.2016.08.005).

[83] Schultz W. 2015. Neuronal reward and decision signals: from theories to data. *Physiol. Rev.* **95**:853–951. (doi:10.1152/physrev.00023.2014).
[84] 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).
[85] 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).
[86] 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).
[87] 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).
[88] 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).

## Supplementary Figures

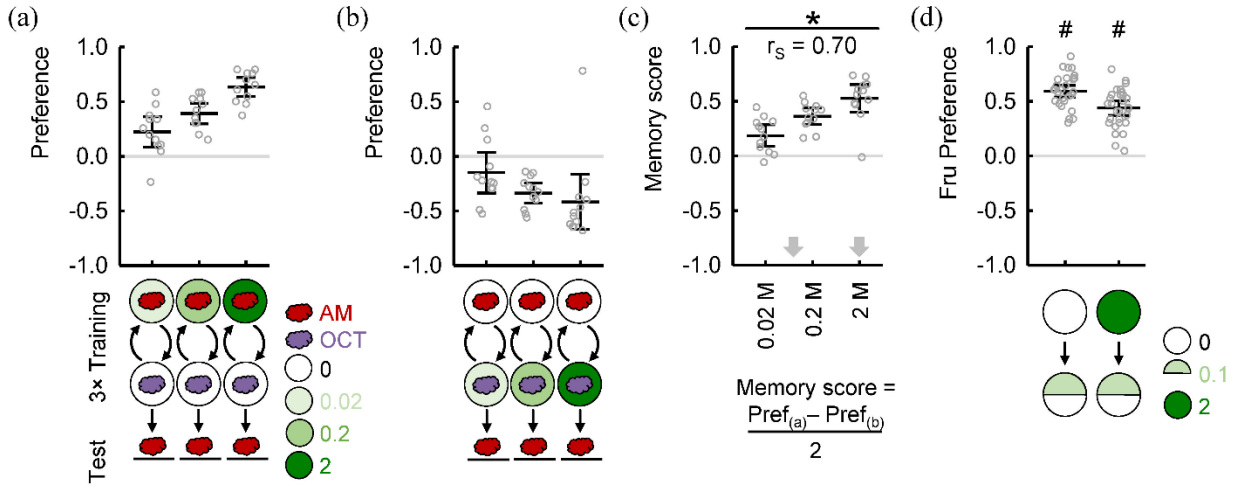


**Figure S1.** Reward strength increases with increasing fructose concentration. Larvae were trained with two odours, AM and OCT, and different concentrations of fructose (0.02, 0.2 and 2 M). (a) When AM had been paired with the fructose reward, the AM preference increased with increasing fructose concentration. (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 S2. (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). For all raw data along with the detailed results of the statistical tests, see supplementary table S1.

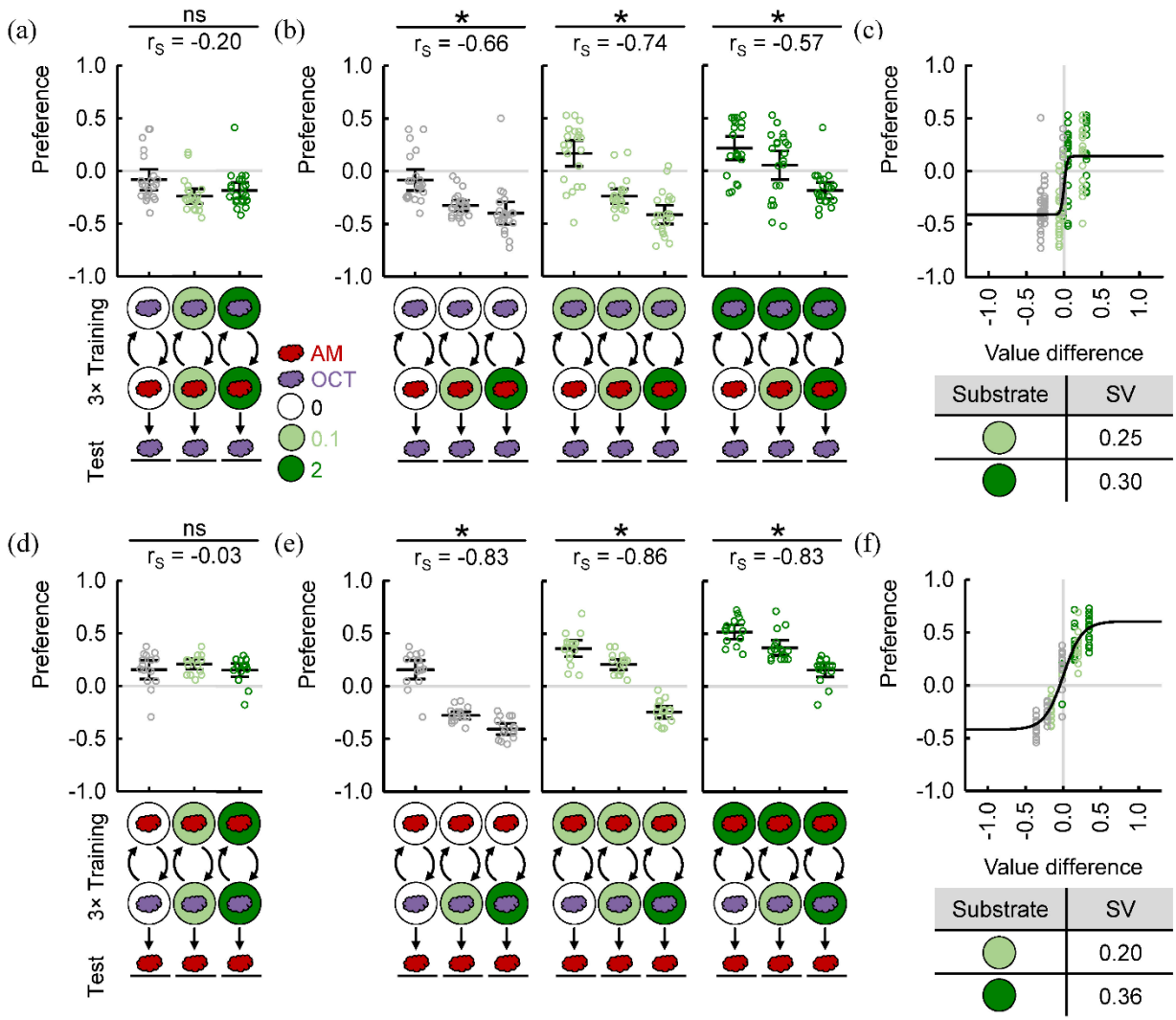


**Figure S2.** Larvae can learn the relative value of rewards independent of the tested odour or the test condition. (a-c) Larvae were trained with two odours, AM and OCT, and different concentrations of fructose. Afterwards, they were tested for their preference for OCT. (a) When the same fructose concentration 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. For all raw data along with the detailed results of the statistical tests, see supplementary table S1.

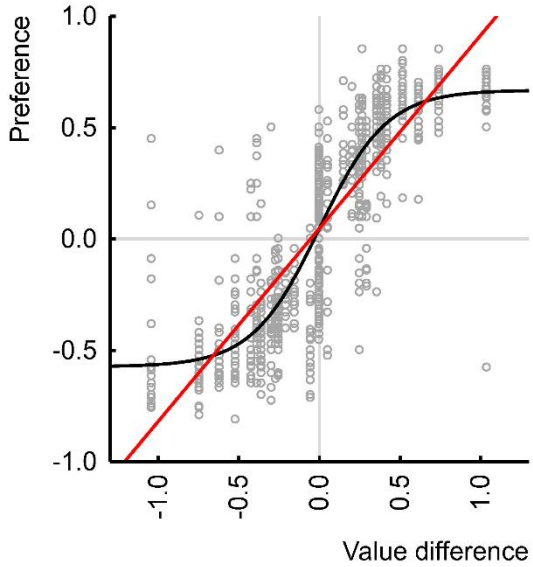


**Figure S3.** The relationship between the value difference and the odour preference is logistic. This graph displays all data of this study using 3 training trials (Fig. 1, 2, S2). To compare a generalized logistic function with a simple linear function, we calculated the Akaike information criterion (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.

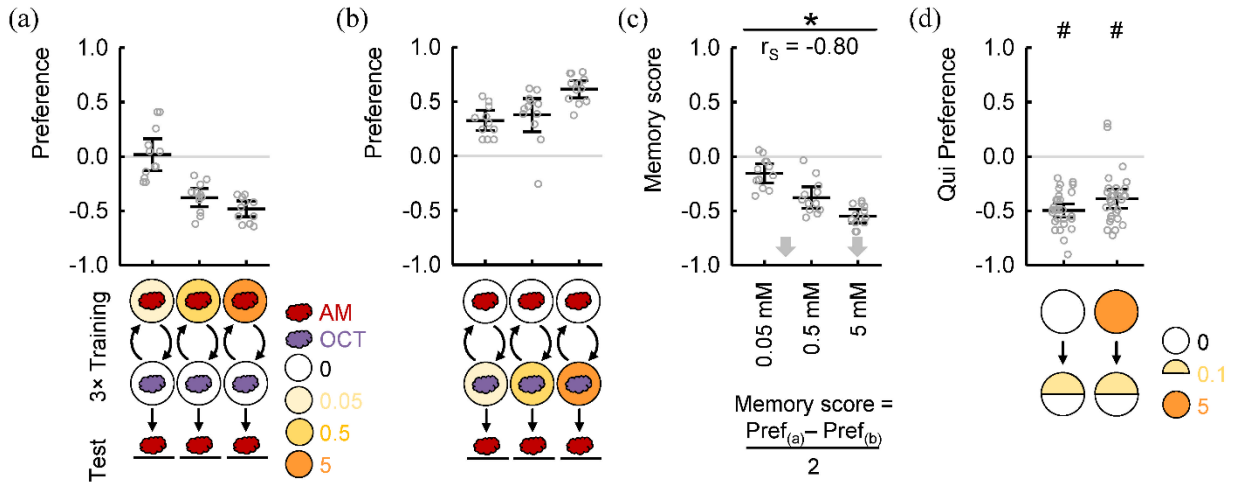


**Figure S4.** Punishment strength increases with increasing quinine concentration. Larvae were trained with two odours, AM and OCT, and different concentrations of quinine (0.05, 0.5 and 5 mM). (a) When AM had been paired with the quinine punishment, the AM preference decreased with increasing quinine 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). For all raw data along with the detailed results of the statistical tests, see supplementary table S1.