



OPEN

Goal-directed behavior in *Tenebrio molitor* larvae

Andrea Dissegna^{1,2}, Massimo Turatto² & Cinzia Chiandetti¹✉

Can signs of intentional behavior be traced in an insect larva, traditionally thought to be driven only by mere reflexes? We trained *Tenebrio molitor* coleoptera larvae in a uniform Y-maze to prefer one target branch to get access to food, observing their ability to learn and retain access to the reward-associated side for up to 24 h. During reward devaluation, the reward food (experimental group) and a different food (control group) were paired with an aversive stimulus in a new environment. When tested again in the Y-maze, mealworms of the experimental group significantly reduced their visits to the target branch, whereas mealworms of the control group did not. Importantly, we found that the larvae did not have to experience the unpleasant consequences directly in the target branch to halt their behavior, as the exposure to the aversive taste occurred in a separate unfamiliar context. This is evidence that the mealworms formed a mental representation of action-consequence relationships, demonstrating flexible control of their actions to achieve desired outcomes at an early stage of their development.

Keywords Conditioned place preference, Outcome devaluation, Conditioning, Intentionality, Invertebrates

Descartes believed that humans were the only creatures capable of thinking and intentional behavior, whereas other animals were merely seen as sophisticated automata lacking mental life and behaving mechanistically according to fixed routines^{1,2}. Intentional behavior differs from fixed routines like habits and reflexes based on two empirical criteria^{3,4}: (1) the instrumental criterion, which requires an understanding of the instrumental contingency between the action and its outcome⁵⁻⁷; and (2) the goal criterion, which implies a representation of the outcome as a goal for the agent^{3,8-10}. Indeed, only goal-directed actions demand that animals possess knowledge of their consequences and the ability to assess the outcome value. Consequently, these actions must be modified or omitted when they have no causal role in obtaining the reward, or when the value of the expected outcome changes¹¹.

Goal-directed behavior is traditionally tested by means of the outcome-devaluation paradigm initially developed with rats¹². Animals are trained through operant conditioning to perform a specific instrumental response, such as pressing a lever, to obtain a reward (i.e. the desired food). Subsequently, the same food is devalued by pairing it with an aversive experience, which makes the food no longer desirable¹³. Note that during the devaluation treatment, the animals do not perform the instrumental response, thus preventing any link between the response and the unpleasant state. Crucially, when later given the chance to press the lever again, rats demonstrate reluctance to do so. This is taken as evidence that the animals have formed a mental representation of the causal relationship between their response and the outcome, and, as a result, they do not perform the action that would lead to the unpleasant consequence, the now devalued food. Symmetrically, this also demonstrates that during the initial instrumental training the animals behave to attain their goal, the originally desired food (e.g. ^{11,12}). In human participants, both adults^{14,15} and infants as young as 2 years of age^{16,17} consistently respond by choosing the optimal alternative after devaluation, independently from verbal competencies.

While there is evidence that mammals like humans and rats can flexibly inhibit instrumental responses leading to devalued outcomes by updating the value of their goals after devaluation treatment, this capacity has not yet been demonstrated in invertebrates.

Importantly, the inhibition of learned responses achieved with the outcome-devaluation paradigm is different from that observed with other paradigms used to research associative learning in invertebrates, including aversive learning¹⁸⁻²¹, reversal learning²²⁻²⁶, and second-order aversive conditioning^{27,28}. The key distinction is that aversive and reversal learning involve the direct suppression of responses that are contingent upon negative outcomes, such as toxins, other aversive substances, and, in the case of second-order aversive conditioning, stimuli associated with negative outcomes. For example, Lai et al.²⁹ trained bees to associate an odorant (conditioned stimulus) with an appetitive fructose solution (unconditioned stimulus), which elicited a Proboscis Extension

¹Department of Life Sciences, University of Trieste, Trieste, Italy. ²CIMeC - Center for Mind/Brain Sciences, University of Trento, Trento, Italy. ✉email: cchiandetti@units.it

Reflex (PER) through Pavlovian conditioning. In the second phase of their experiment, the authors exposed the bees to a conditioned taste aversion, where the PER elicited by fructose led to the ingestion of quinine, causing malaise in the bees. In a subsequent test, Lai and colleagues observed that the bees reduced their PER response to both the fructose and the odorant, indicating that the bees remembered the negative outcome of this response from the previous phase and decreased the strength of the association between PER, fructose, and the odorant.

Conversely, in the outcome-devaluation paradigm, a reward's value is reduced independently of the associated instrumental response^{3,12}. This paradigm, indeed, capitalizes on the fact that animals must rely on associative memories (i.e., the knowledge of the instrumental contingencies between actions and their outcome, and the representation of the reward as the final goal of an acquired response) to suppress behaviors leading to devalued goals, without needing to directly experience the now-negative outcomes of those behaviors.

To demonstrate this ability in invertebrates, we focused on larvae of *Tenebrio molitor*. We chose this species for two reasons. First, few studies focused on the learning capacity of this insect larvae, compared to other species^{23,30–32} even though adult individuals possess rudimentary numerical abilities³³, and the pupae are already capable of habituation³⁴. Furthermore, the larva shows spatial learning³⁵, memory retention³⁶, also across the metamorphosis from larval to adult stage^{35,37,38}. Second, *T. molitor* larvae have been recently considered as edible food by FAO³⁹ therefore urging researchers to advance the knowledge on their abilities to improve their protection and welfare.

Results

We adopted a slightly modified version of the conditioned place preference procedure, which is based on the learned preference for one place over a different one, generated after pairing it with the rewarding stimuli^{40,41}. This learned behavioral response is observed in various invertebrates^{42–46}. It involves the animal showing a preference for a specific location that has been previously associated with rewarding events. In our case, it consisted with the location of the rewarded branch (left or right) of a Y-maze paired with an apple slice (Fig. 1A). Subsequently, animals were submitted to the reward devaluation phase, where for the experimental group, the reward was paired with an aversive stimulus (a lemon drop, Fig. 1B), whereas for the control group, we used a different food (wheat bran), thus leaving the original food reward used during training unaffected by the devaluation procedure (see the Supplementary Information for details). Finally, in the last test phase, the animals were re-tested in the Y-maze task under extinction to evaluate their willingness to choose the target branch where they received the food reward during the initial training (Fig. 1C). However, after the devaluation phase this outcome would have been no longer desirable for the animals of the experimental group, whereas it should have been still palatable for the animals of the control group. If the choice of the Y-maze branch is a goal-directed act guided by the belief that it will lead to a given outcome, we expected the experimental group to reduce this behavior in the test phase compared to the training phase, whereas we predicted no difference between the two phases in the control group, as for this group the reward used in the training phase was never devalued.

The results showed that the by-session average proportion of visits to the apple-associated branch was at chance level during the first conditioning block (60%, SE = 11.6%, difference from the 50% chance level: $z = 1.59$, $p = 0.110$), but it increased significantly during the last block of the day (85%, SE = 8%, difference from chance level: $z = 3.19$, $p < 0.001$). All 27 mealworms met the 80% criterion at the end of the first day (Fig. 1D). On average, a mealworm had to conclude 2.33 training blocks before reaching this criterion, with variations across individuals in the min–max range of 2–4 blocks (standard error, s.e. = 0.12). The next day, in the “pre-devaluation” test, 20 of the 27 mealworms (74%) still met the criterion (84%, SE = 8.2%, difference from chance level: $z = 3.48$, $p < 0.001$), indicating long-term memory of the conditioned place preference for the apple-associated branch. The long-term retention of the place preference was further corroborated by the linear model used to assess the proportion of visits to the target branch (its estimates are reported in Table 1), which revealed an overall significant increment of the visits ($F(2, 295) = 8.23$, $p < 0.001$, $\eta_p^2 = 0.64$).

We randomly assigned the mealworms meeting the 80% criterion to either an experimental or a control group (see the “Materials and methods” section, $n = 10$ each). We conducted a post-hoc analysis on the proportion of visits to the target branch including the treatment group as an additional factor of a linear model to test for differences between the two groups during the conditioned place preference phase. The model revealed a main effect of the testing session ($F(2, 283) = 10.94$, $p < 0.001$, $\eta_p^2 = 0.65$), consistent with an overall increment of the visits to the target branch, but no effect of the group and interaction between group and testing session (all $ps > 0.050$), supporting that the two groups were similar before the outcome-devaluation procedure.

Crucially, after apple devaluation (Fig. 1E), the animals in the experimental group stopped visiting the target branch from the first trial post-devaluation (0%, SE = 14.2%; average visits in the “post-devaluation” phase for this group = 14.5%, SE = 10.7%). By contrast, animals in the control group, initially continued to visit the target branch consistently with their pre-devaluation behavior (100%, SE = 10.5%), then gradually decreased their visits across trials (average visits: 58%, SE = 15.8%), consistent with the fact that this behavior was no longer reinforced by the food reward and was consequently extinguished.

This pattern of results was supported by the model comparing the visits to the target branches during the 5th trial “pre-devaluation” with the 1st trial “post devaluation”, which revealed an effect of the trial ($F(1, 33) = 39.23$, $p < 0.001$, $\eta_p^2 = 0.87$), of the group ($F(1, 33) = 38.69$, $p < 0.001$, $\eta_p^2 = 0.86$), and their interaction ($F(1, 33) = 61.30$, $p < 0.001$, $\eta_p^2 = 0.91$); the model estimates are reported in Table 2). Specifically, the proportion of visits to the target branch decreased from the 5th trial “pre-devaluation” to the 1st trial “post-devaluation” in the Experimental group ($\chi^2(1) = 99.31$, *Holm's* $p < 0.001$), but not in the Control group ($\chi^2(1) = 1.22$, *Holm's* $p = 0.268$), indicating a specific devaluation effect. We also analyzed the proportion of mealworms' visits to the target branches during the “post-devaluation” test. In this case, the model revealed a main effect of the group ($F(1, 85) = 8.90$, $p < 0.001$, $\eta_p^2 = 0.39$), and a significant trial \times group interaction ($F(4, 85) = 6.34$, $p < 0.001$, $\eta_p^2 = 0.64$). Estimates

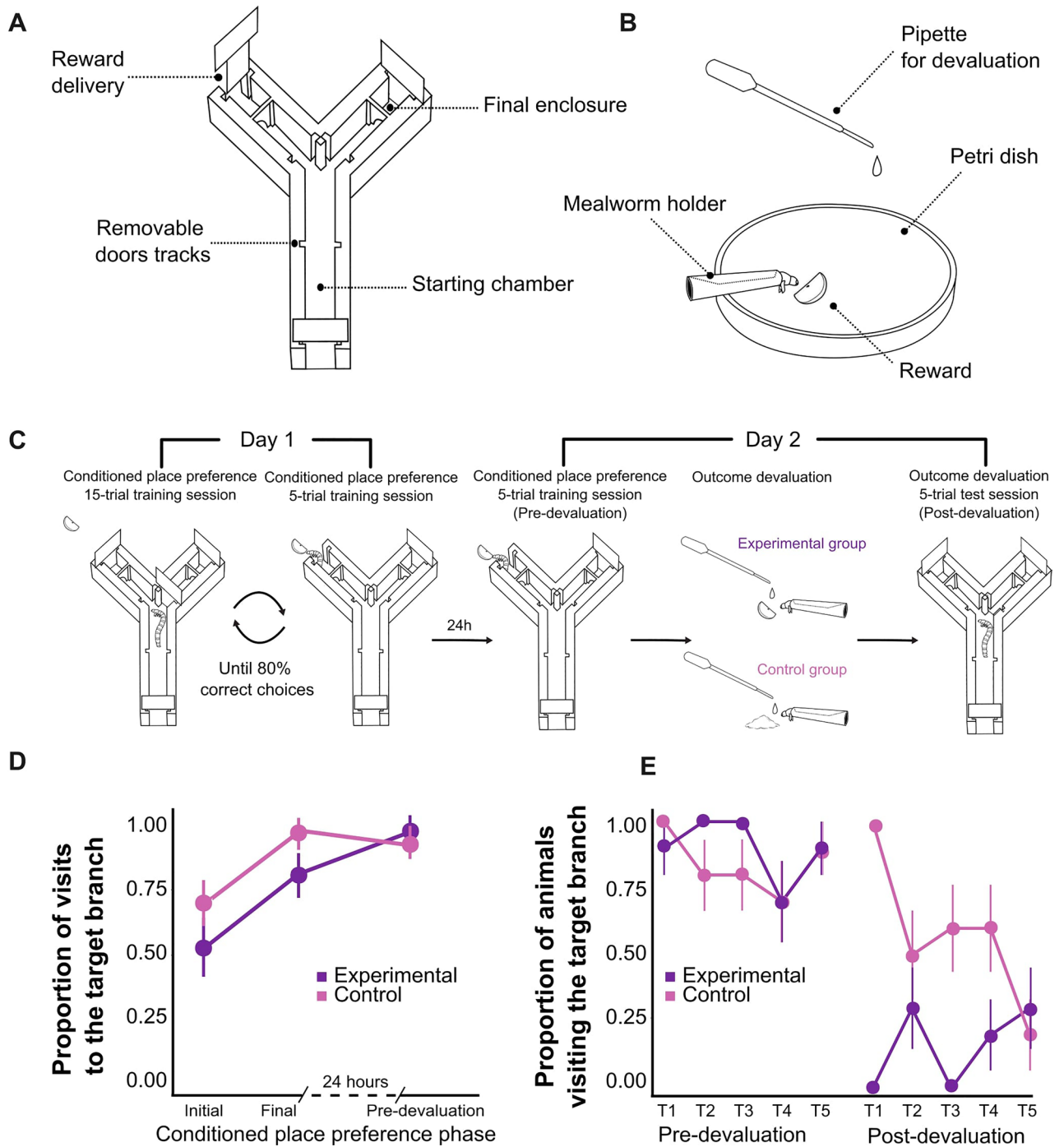


Fig. 1. Goal-directed behavior in *Tenebrio molitor* larvae. **(A,B)** Setups for conditioned place preference and outcome devaluation. **(A)** Mealworms learned the position of a rewarded branch (left or right) of a Y-maze, using an apple slice as reward. The conditioned place preference involved a continuous alternation between two distinct trial sessions: the training session and the subsequent 5-trial sessions (refer to Supporting information for further details). The alternation of these two sessions continued until worms achieved a set criterion of visits to the target branch, corresponding to 80% of the trials (i.e., 4 out of 5 trials) of a 5-trial session. **(B)** The reward used for conditioning the place preference was devalued in the Experimental group. The devaluation was accomplished by presenting the worms with an aversive stimulus, namely lemon, immediately after their consumption of an apple slice. Another food (wheat bran) was devalued in the Control group, following an identical procedure. This phase took place within a petri dish to maximize the difference between the devaluation site and the testing apparatus. **(C)** Overview of the protocol used during the conditioning place preference and the outcome-devaluation phase. **(D,E)** Results of the experiment. **(D)** By-session average proportion of visits to the target branch across the conditioned place preference phases in the Experimental (violet line) and the Control (pink line) group. **(E)** Proportion of visits to the target branch by trial and treatment groups (the violet line represents the Experimental group and the pink line represents the Control group) before and after devaluation of the reward.

Variable	Estimate	Std. error	z value	Pr(> z)
Initial session (intercept)	0.29	0.18	1.6	0.110
Final session	0.75	0.23	3.19	<0.001
Pre-devaluation	0.78	0.22	3.48	<0.001

Table 1. Estimates of the model (with “probit” link) used to analyze the visits to the target branch during the initial and final testing session of the first day and on “pre-devaluation” test, 24 h later. Estimates refer to the by-session average proportion of visits to the target branch.

Variable	Estimate	Std. error	t value	Pr(> t)
5th trial experimental group (intercept)	0.9	0.07	12.4	<0.001
1st trial experimental group	−0.9	0.09	−9.97	<0.001
5th trial control group	0	0.1	0	1
1st trial control group	1	0.13	7.83	<0.001

Table 2. Estimates of the model (with “probit” link) used to analyze the visits to the target branch on the 5th trial “pre-devaluation” and on the 1st trial “post-devaluation” in each condition. Estimates and standard errors derives from a Bayesian linear mixed-effects model (see the “Materials and methods” section).

of this model are reported in Table 3. Specifically, the proportion of visits to the target branch did not change across “post-devaluation” trials in the experimental group ($\chi^2(1) = 3.31$, *Holm's* $p = 0.959$), whereas it decreased for mealworms in the control group ($\chi^2(1) = 23.59$, *Holm's* $p < 0.001$).

Discussion

Conditioned place preference is widely used to study the motivational effects of natural and pharmacological stimuli, especially drugs of addiction⁴¹. Pavlovian conditioning is believed to play a key role in this behavioral phenomenon, as the unconditioned response triggered by the rewarding event (the unconditioned stimulus) would become associated with some sensory (e.g., visual, tactile) properties of the place where the reward is experienced. Therefore, the reward-associated place would function as a form of conditioned stimulus capable of triggering a Pavlovian conditioned approach. This ends in the heightened frequency of visits to the reward-associated compartment over the unrewarded-compartment^{47–50}. It is important to highlight that for this form of Pavlovian conditioning to take effect in the conditioned place paradigm, the location where the reward is delivered is made perceptually distinct from the other, which was not the case of our apparatus. Without this distinction, the observed place preference would hardly occur. Moreover, while Pavlovian conditioning can certainly be an important factor, it has been argued that other forms of learning, including operant conditioning, can contribute to the conditioned place preference phenomenon⁴⁰. In line with this possibility, there are compelling reasons to believe that, in our paradigm, the branch preference emerged because of operant conditioning rather than classical conditioning. Hence, the larvae's choice to visit the target branch would be more like an arbitrary instrumental response rather than an unconditioned approach response. This perspective is supported by two main considerations. Firstly, in contrast to the conventional conditioned place preference protocol, where the drug reward is administered before placing the animal in the designated compartment, our paradigm delayed the reward delivery until the animal reached the end of the target branch. In this way, the reward cannot trigger an unconditioned approach response that might otherwise be associated with the visual cue of a specific location. Secondly, any incentive-motivational value attributed to the food reward is improbable to have been transferred, through Pavlovian mechanisms, to the defining feature of the target branch, since both branches of the Y-maze were identical. On the contrary, a more direct and straightforward explanation for our results is that the larvae

Variable	Estimate	Std. error	t value	Pr(> t)
Experimental group (intercept)	0.01	0.13	0.08	0.939
Trial	0.05	0.04	1.31	0.200
Control group	1.02	0.22	4.74	<0.001
Trial: control group	−0.2	0.05	−3.69	<0.001

Table 3. Estimates of the model (with “probit” link) used to analyze the visits to the target branch during the “post-devaluation” test in each condition. Estimates and standard errors derives from a bayesian linear mixed-effects model. The estimates of “Experimental group” and “Control group” refers to the intercepts of the best fitting regression lines estimated by the linear model for the two treatment groups; the estimates of “Trial” and “Trial:Control group” refers to their respective slopes (see the “Materials and methods” section).

learned to select the correct target branch due to this behavior being consistently reinforced by the food reward provided when they reached the end of the branch. In this context, the larvae's choice of the target branch parallels the lever press used by rats to obtain a food pellet in analogous experiments¹². Hence, if it is accepted that rats exhibit goal-directed behavior by refraining from pressing the lever after reward devaluation, then the same holds true for *Tenebrio molitor* larvae when they refrain from visiting the target branch after reward devaluation.

However, if visiting the target branch was solely a conditioned Pavlovian approach response, our results would be even more remarkable. Indeed, they would demonstrate that even unconditioned responses are under goal-directed control, which is considered a hallmark of intentionality^{3,8,9}. Of course, our findings do not imply that all of the larva's behavior is necessarily intentional, as in the case of human beings, whose actions are also influenced by habits and reflexes^{14,15}. Importantly, our results underscore the complexity of the mental life of an invertebrate species, traditionally considered as an *automata*, a creature that would not be capable of voluntary action, but that merely responds to external stimuli by means of reflexes⁵¹.

An alternative interpretation of our results would assume that the devaluation procedure did not specifically target the devalued reward, but instead led to a general decrease in food interest, which would challenge our conclusion that mealworms intentionally reduced their visits to the branch associated with the devalued food. However, our pilot experiment 2 (Supplementary Information) ruled out this possibility by demonstrating that after devaluing a given reward, a different reward was still approached and considered palatable by the animals.

Showing goal-directed behavior in the *Tenebrio molitor* larvae is a fundamental first step towards studying intentionality in this species. Our results indicate that these animals in early stage of their life already possess a higher level of cognitive abilities than intuitively and previously assumed. They also reveal that these insect larvae are capable of complex decision-making processes, including evaluating different options, weighing costs and benefits, and selecting actions based on their internal goals and motivations. Hence, these insect larvae can actively assess and respond to changing environmental conditions, demonstrating flexibility to achieve the desired outcomes. This adaptability parallels that of other insects, such as bumblebees, who rely on associative memories, rather than direct experience of the stimuli, to control their behavior⁵².

Our study has also ethical implications, because it highlights the existence of a complex and rich mind in these simple animals, emphasizing that they are capable of purposeful actions. This can influence humans' attitudes and behaviors even towards very simple creatures, like mealworm studied here, probably the most popular larva used either to feed reptiles, birds, poultry, or as fishing bites, and more recently suggested for a protein-rich diet for humans. Overall, demonstrating goal-directed behaviors in insects adds to the evidence of complexity in their cognitive ability, and likely promoting empathy, and potentially impacting animal welfare and conservation efforts.

Materials and methods

Participants

T. molitor at the larval stage were sourced from a local animal shop and cultured in our laboratory. The mealworms were housed in a plastic box measuring 20 × 36 × 12 cm, equipped with a fine-mesh nylon lid. The colony was maintained under a 12:12 light cycle at conditions of 26 °C with 50% relative humidity and fed with wheat bran. From the main colony, we selected mealworms at the last larval stages, i.e., with a body length ranging between 2.5 and 3.5 cm (from spine to head). Two days before the experiment, the selected mealworms were transferred to plastic boxes measuring 10 × 12 × 6 cm and maintained at identical environmental conditions as those recorded for the main colony. In this period, mealworms were food-deprived to increase their foraging motivation during the experiments. After the end of the experiments, all mealworms completed their life cycle and all the procedures we adopted were intended to keep at minimum their stress and avoid pain. Mealworms are not included in the European law on animal research, nonetheless in this experiment we minimized the number of animals used. We tested a total of 97 mealworms across two pilot experiments (reported in the Supplementary Information) and the main experiment reported in this manuscript. The mealworms tested in the main experiment were part of a naïve control condition of the first pilot experiment that was not exposed to any reward or aversive substance, to capitalize on individuals who were already familiar with the apparatus and whose spontaneous behavior in the apparatus was known. None of the selected mealworms molted during the experiments.

Experimental apparatus

The apparatus consisted of a 3D-printed Y-maze made of white PLA*, that ensured color and surface uniformity of the walls within the maze. The Y-maze consisted of various sections connected by removable doors. The Y-stem included a starting chamber and a hallway, both measuring 3 × 3 × 1 cm (height, length, and width). The angle between the Y-branches was 45°, measured 3 × 3 × 1 cm, and included a separated enclosure at their end, measuring 3 × 1 × 1 cm, with a semi-circular opening of radius 1 cm. Exit from the maze was allowed through a removable door, located on the opposite side of this enclosure. The reward was delivered at the end of rewarded trials on a food tray, behind this door, and was not visible to mealworms.

Apple and wheat bran were used as rewards, while lemon was used as “devaluator”. We opted for lemon as we observed that mealworms rapidly developed an aversion for its aversive taste.

Procedure

We devised two pilot experiments: one to control for any spontaneous preference for one branch of the maze, and the other to control for the impact of the devaluation procedure on the general feeding motivation of mealworms. Details of the pilot experiments are reported as Supplementary Information.

The experiment consisted of two main phases: one involved conditioning a preference for a target (left/right) branch of the Y-maze using apple as a reward (i.e., the conditioned place preference phase), and the other involved

giving to the same mealworms either apple or wheat bran, and then a drop of lemon as an aversive stimulus (i.e., the outcome devaluation phase).

Conditioned place preference phase

Individuals were placed for 2 min in the starting chamber of the maze to familiarize with the new environment. After this time, the door of the starting chamber was removed allowing the mealworms to enter the hallway of the maze. When a mealworm reached the final enclosure of the target branch, its behavior was rewarded by giving access for 30 s to an apple slice (1 g); in contrast, reaching the end of the other branch led no consequence for the mealworm. At the end of each trial, the mealworm was placed back in the starting chamber, and a new trial began after 1 min.

Mealworms were trained to associate the target branch with the apple reward over 3 training sessions occurring on the same day. Each session consisted of 15 trials, with 10 trials in which mealworms could access exclusively the apple-associated branch and 5 trials in which mealworms could access exclusively the unrewarded branch (the sequence of rewarded and unrewarded trials is reported in the Supplementary Information). This feeding ratio was used to maintain the animals' motivation for the food reward. We used the removable doors of the maze to control which branch was available to mealworms on each trial. The target branch was counterbalanced across mealworms.

Each training session was followed by a shorter 5-trial session, in which mealworms could enter both branches of the maze and visits of the target branch were rewarded with apple. An additional 5-trial session was conducted before starting the outcome-devaluation phase the next day (i.e., the "pre-devaluation" test) to assess long-term retention of the conditioned place preference.

Outcome-devaluation phase

Mealworms that visited the target branch at least 80% of the trials during the "pre-devaluation" test were considered to have successfully associated the target branch with apple and were passed to the outcome-devaluation procedure. In this phase, animals were individually confined within specific funnel-shaped holders made of paper (see Fig. 1B) and presented with an appetitive stimulus consisting of an apple slice for the "experimental group", and wheat bran for the "control group". Each mealworm was allowed to consume the food for 30 s before being given a 10 μ l of lemon, administered by bringing a pipette in contact with its mouth, and serving as an aversive stimulus. This aversive conditioning procedure was repeated until a mealworm retreated in the holder for a period of 3 min, indicating an aversion for the food, which usually took place after 2 exposures to lemon (min = 2; max = 3). It should be noted that, for both groups, food was paired with lemon outside the Y-maze, thus making impossible for the animals to associate the aversive conditioning with a given Y-maze branch.

Next, all mealworms underwent a 5-trials "post-devaluation" session in the Y-maze to test any change in the preferred branch compared to the initial conditioned place preference; in this phase, their behavior was never rewarded.

Mealworms' performance was recorded by a Logitech® camera, and video were off-line analyzed. Visits to the target branch were scored 1, whereas visits to the other branch and trials in which a mealworm did not reach the final enclosure of a branch within 5 min were scored 0. The inter-trial interval in this phase was 1 min, during which mealworms remained in the starting chamber of the Y-maze.

Data analysis

Conditioned place preference phase. We analyzed the proportion of visits to the apple-associated branch in each 5-trial session. We considered: (1) the number of mealworms reaching the 80% visits criterion at the end of the first day; (2) the number of training sessions before reaching this criterion; (3) the average proportion of visits to target branch on the initial and the final 5-trial session of the first day and on the "pre-devaluation" test, the next day. We used generalized mixed linear models (*probit* link) with subject-specific slope and intercept, with the test phase ("initial", "final", and "pre-devaluation") as factor. Models were fitted using the `glmer()` function of the `lmerTest` package⁵³. We conducted a sensitivity analysis with $\alpha = 0.05$ and power set to 0.95 to determine the minimal detectable effect for our experimental design. We found that this effect was $\eta_p^2 = 0.08$, corresponding to $F = 3.24$.

Outcome-devaluation phase. We assessed whether the preference for the apple-associated branch changed after the devaluation procedure. We thus compared the proportion of visits to the target branch during the 5th trial of the "pre-devaluation" session with the 1st trial of the "post-devaluation" session.

As shown in the "Results" section, in the 1st post-devaluation trial, all mealworms of the experimental group avoided the target branch, and this led to a "complete separation problem" making impossible to reliably use standard generalized linear models due to infinite variance estimates. To overcome this limitation, we used Bayesian generalized mixed linear models. These models blend Bayesian regularization with specific low informative priors, counteracting the risks of unbounded or skewed estimates⁵⁴. Within this Bayesian framework, we maintained the assumption of a normal distribution for fixed effects (similar to the *probit* link used in standard generalized models used in the previous experiments) and fitted the data using the `bgfmlmer()` function (from the `blme` package⁵⁵), keeping all the settings as recommended by the package authors. Our model included the effects of session ("pre-devaluation" and "post-devaluation"), group ("Experimental" vs. "Control"), and their interaction, in both fixed and random effects structures. A sensitivity analysis, set with $\alpha = 0.05$ and power at 0.95, revealed that the minimal detectable effect for this analysis was $\eta_p^2 = 0.09$, corresponding to $F = 4.41$.

In addition, we used a similar model to analyze mealworms' visits to the target branch throughout the trials of the "post devaluation" session. To this aim, we used a model incorporating the effects of trial (from "trial 1"

to “trial 5”, considered as a continuous variable), group (“experimental” vs. “control”), and their interaction in its fixed and random effects structure. We applied Holm’s correction of p -values resulting from post-hoc analysis of the interactions. A sensitivity analysis (same α and power as before) revealed a minimal detectable effect of $\eta_p^2 = 0.05$ ($F = 2.49$).

Data availability

Data is provided within the Supplementary Information files.

Received: 11 July 2024; Accepted: 6 September 2024

Published online: 17 September 2024

References

1. Descartes Discourse on the method. In *The Philosophical Writings of Descartes* Vol. 1 (eds Cottingham, J. et al.) 111–176 (Cambridge University Press, 2020).
2. Chiandetti, C. *How the Evolutive Continuity of Cognition Challenges ‘Us/Them’ Dichotomies* (EUT Edizioni Università di Trieste, 2020).
3. Dickinson, A. & Balleine, B. Motivational control of goal-directed action. *Anim. Learn. Behav.* **22**, 1–18 (1994).
4. Dickinson, A. Intentionality in animal conditioning. *Thought Language* 305–325 (1988).
5. Dickinson, A., Watt, A. & Griffiths, W. J. H. Free-operant acquisition with delayed reinforcement. *Q. J. Exp. Psychol. Sect. B* **45**, 241–258 (1992).
6. Hammond, L. J. The effect of contingency upon the appetitive conditioning of free-operant behavior. *J. Exp. Anal. Behav.* **34**, 297–304 (1980).
7. Colwill, R. M. & Rescorla, R. A. Associative structures in instrumental learning. In *Psychology of Learning and Motivation* Vol. 20 (ed. Bower, G. H.) 55–104 (Academic Press, 1986).
8. Gómez, J.-C. Intentionality. In *Encyclopedia of Animal Cognition and Behavior* (eds Vonk, J. & Shackelford, T. K.) 3614–3622 (Springer, 2022).
9. Dennett, D. C. Intentional systems in cognitive ethology: The “Panglossian paradigm” defended. *Behav. Brain Sci.* **6**, 343–355 (1983).
10. Colwill, R. M. & Rescorla, R. A. Postconditioning devaluation of a reinforcer affects instrumental responding. *J. Exp. Psychol. Anim. Behav. Process.* **11**, 120–132 (1985).
11. Heyes, C. & Dickinson, A. The intentionality of animal action. *Mind Lang.* **5**, 87–103 (1990).
12. Adams, C. D. & Dickinson, A. Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol. Sect. B* **33**, 109–121 (1981).
13. Garcia, J., Kimeldorf, D. J. & Koelling, R. A. Conditioned aversion to saccharin resulting from exposure to gamma radiation. *Science* **122**, 157–158 (1955).
14. de Wit, S., Niry, D., Wariyar, R., Aitken, M. R. F. & Dickinson, A. Stimulus-outcome interactions during instrumental discrimination learning by rats and humans. *J. Exp. Psychol. Anim. Behav. Process.* **33**, 1–11 (2007).
15. Valentin, V. V., Dickinson, A. & O’Doherty, J. P. Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* **27**, 4019–4026 (2007).
16. Klossek, U. M. H., Russell, J. & Dickinson, A. The control of instrumental action following outcome devaluation in young children aged between 1 and 4 years. *J. Exp. Psychol. Gen.* **137**, 39–51 (2008).
17. Kenward, B., Folke, S., Holmberg, J., Johansson, A. & Gredebäck, G. Goal directedness and decision making in infants. *Dev. Psychol.* **45**, 809–819 (2009).
18. Dyer, A. G. & Howard, S. R. Aversive reinforcement improves visual discrimination learning in free-flying wasps (*Vespa vulgaris*). *Behav. Ecol. Sociobiol.* **77**, 101 (2023).
19. Jelen, M., Musso, P.-Y., Junca, P. & Gordon, M. D. Optogenetic induction of appetitive and aversive taste memories in *Drosophila*. *eLife* **12**, e81535 (2023).
20. Dwijesh, A. S., Eswaran, A., Berry, J. A. & Phan, A. Diverse memory paradigms in *Drosophila* reveal diverse neural mechanisms. *Learn. Mem.* **31**, a053810 (2024).
21. Klappenbach, M., Lara, A. E. & Locatelli, F. F. Honey bees can store and retrieve independent memory traces after complex experiences that combine appetitive and aversive associations. *J. Exp. Biol.* **225**, 244229 (2022).
22. Behmer, S. T., Belt, C. E. & Shapiro, M. S. Variable rewards and discrimination ability in an insect herbivore: What and how does a hungry locust learn? *J. Exp. Biol.* **208**, 3463–3473 (2005).
23. Mancini, N. et al. Reversal learning in *Drosophila* larvae. *Learn. Mem.* **26**, 424–435 (2019).
24. Matsumoto, Y. & Mizunami, M. Olfactory learning in the cricket *Gryllus bimaculatus*. *J. Exp. Biol.* **203**, 2581–2588 (2000).
25. McCurdy, L. Y., Sareen, P., Davoudian, P. A. & Nitabach, M. N. Dopaminergic mechanism underlying reward-encoding of punishment omission during reversal learning in *Drosophila*. *Nat. Commun.* **12**, 1115 (2021).
26. Young, F. J., Melo-Flórez, L., McMillan, W. O. & Montgomery, S. H. Reversal learning of visual cues in Heliconiini butterflies. *Anim. Behav.* **208**, 69–77 (2024).
27. Tabone, C. J. & de Belle, J. S. Second-order conditioning in *Drosophila*. *Learn. Mem.* **18**, 250–253 (2011).
28. Mizunami, M. et al. Roles of octopaminergic and dopaminergic neurons in appetitive and aversive memory recall in an insect. *BMC Biol.* **7**, 46 (2009).
29. Lai, Y. et al. Degradation of an appetitive olfactory memory via devaluation of sugar reward is mediated by 5-HT signaling in the honey bee. *Neurobiol. Learn. Mem.* **173**, 107278 (2020).
30. Salloum, A., Colson, V. & Marion-Poll, F. Appetitive and aversive learning in *Spodoptera littoralis* larvae. *Chem. Senses* **36**, 725–731 (2011).
31. Schroll, C. et al. Light-induced activation of distinct modulatory neurons triggers appetitive or aversive learning in *Drosophila* larvae. *Curr. Biol.* **16**, 1741–1747 (2006).
32. Dacks, A. M., Riffell, J. A., Martin, J. P., Gage, S. L. & Nighorn, A. J. Olfactory modulation by dopamine in the context of aversive learning. *J. Neurophysiol.* **108**, 539–550 (2012).
33. Carazo, P., Fernández-Perea, R. & Font, E. Quantity estimation based on numerical cues in the mealworm beetle (*Tenebrio molitor*). *Front. Psychol.* **3**, 2 (2012).
34. Bernal-Gamboa, R., García-Salazar, J. & Gámez, A. M. Analysis of habituation learning in mealworm pupae (*Tenebrio molitor*). *Front. Psychol.* **12**, 1 (2021).
35. Borsellino, A., Pierantoni, R. & Schieti-Cavazza, B. Survival in adult mealworm beetles (*Tenebrio molitor*) of learning acquired at the larval stage. *Nature* **225**, 963–964 (1970).
36. Walrath, L. C. Retention and interference in the beetle. *Psychon. Sci.* **18**, 267–268 (1970).
37. Alloway, T. M. Retention of learning through metamorphosis in the grain beetle *Tenebrio molitor*. *Am. Zool.* **12**, 471–477 (1972).

38. Alloway, T. M. & Routtenberg, A. ‘Reminiscence’ in the cold flour beetle (*Tenebrio molitor*). *Science* **158**, 1066–1067 (1967).
39. FAO. *Looking at Edible Insects from a Food Safety Perspective* (FAO, 2021).
40. Huston, J. P., de Silva, M. A. S., Topic, B. & Müller, C. P. What’s conditioned in conditioned place preference? *Trends Pharmacol. Sci.* **34**, 162–166 (2013).
41. Tzscentke, T. M. REVIEW ON CPP: Measuring reward with the conditioned place preference (CPP) paradigm: Update of the last decade. *Addict. Biol.* **12**, 227–462 (2007).
42. Baggett, V. *et al.* Place learning overrides innate behaviors in *Drosophila*. *Learn. Mem.* **25**, 122–128 (2018).
43. Zars, M. & Zars, T. Rapid matching in *Drosophila* place learning. *Naturwissenschaften* **96**, 927–931 (2009).
44. Agarwal, M. *et al.* Dopamine and octopamine influence avoidance learning of honey bees in a place preference assay. *PLoS ONE* **6**, e25371 (2011).
45. Jordan, L., Alcalá, J. A., Urcelay, G. P. & Prados, J. Conditioned place avoidance in the planaria *Schmidtea mediterranea*: A pre-clinical invertebrate model of anxiety-related disorders. *Behav. Processes* **210**, 104894 (2023).
46. Crook, R. J. Behavioral and neurophysiological evidence suggests affective pain experience in octopus. *iScience* **24**, 102229 (2021).
47. Bindra, D. A motivational view of learning, performance, and behavior modification. *Psychol. Rev.* **81**, 199–213 (1974).
48. Bolles, R. C. Reinforcement, expectancy, and learning. *Psychol. Rev.* **79**, 394–409 (1972).
49. Spiteri, T., Le Pape, G. & Ágmo, A. What is learned during place preference conditioning? A comparison of food- and morphine-induced reward. *Psychobiology* **28**, 367–382 (2000).
50. Bardo, M. T. & Bevins, R. A. Conditioned place preference: What does it add to our preclinical understanding of drug reward? *Psychopharmacology* **153**, 31–43 (2000).
51. Tolman, E. C. *Purposive Behavior in Animals and Men* (University of California Press, 1932).
52. Gibbons, M., Versace, E., Crump, A., Baran, B. & Chittka, L. Motivational trade-offs and modulation of nociception in bumblebees. *Proc. Natl. Acad. Sci.* **119**, e2205821119 (2022).
53. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: Tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).
54. Chung, Y., Rabe-Hesketh, S., Dorie, V., Gelman, A. & Liu, J. A nondegenerate penalized likelihood estimator for variance parameters in multilevel models. *Psychometrika* **78**, 685–709 (2013).
55. Dorie, V. & Dorie, M. V. *blme: Bayesian Linear Mixed-Effects Models* (2021).

Acknowledgements

We thank Nicolò Petruzzella and Giovanna Sabbatini for help with data collection.

Author contributions

CC and MT conceived the original idea. AD, CC, and MT designed the experiments. AD carried out the experiments, and analyzed the data. CC and MT wrote the original draft. All authors discussed the results and contributed to the final manuscript and revision.

Competing interests

The authors declare no competing interests.

Ethical approval

All the procedures we adopted were intended to keep at minimum their stress and avoid pain. Mealworms are not included in the European law on animal research.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-72455-3>.

Correspondence and requests for materials should be addressed to C.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024