Outcome expectations drive learned behaviour in larval *Drosophila*

Bertram Gerber[*] and Thomas Hendel[†]

Universität Würzburg, Lehrstuhl für Genetik und Neurobiologie, Biozentrum Am Hubland, 970 74 Würzburg, Germany

Why does Pavlov's dog salivate? In response to the tone, or in expectation of food? While in vertebrates behaviour can be driven by expected outcomes, it is unknown whether this is true for non-vertebrates as well. We find that, in the *Drosophila* larva, odour memories are expressed behaviourally only if animals can expect a positive outcome from doing so. The expected outcome of tracking down an odour is determined by comparing the value of the current situation with the value of the memory for that odour. Memory is expressed behaviourally only if the expected outcome is positive. This uncovers a hitherto unrecognized evaluative processing step between an activated memory trace and behaviour control, and argues that learned behaviour reflects the pursuit of its expected outcome. Shown in a system with a simple brain, an apparently cognitive process like representing the expected outcome of behaviour seems to be a basic feature of behaviour control.

Keywords: olfaction; taste; learning; expectation; cognition; outcome

1. INTRODUCTION

While there is no doubt that human behaviour can be guided by the expectation of its outcome, animal behaviour is often thought to be purely stimulus-evoked. In the behaviourist tradition, this extends to learned animal behaviour as well (Rescorla & Wagner 1972; Rescorla 1988). Only relatively recently, experiments on retrospective devaluation in rats (Colwill & Rescorla 1990; Colwill & Motzkin 1994) and causality judgement in man (Dickinson 2001) refuted this dogma in favour of an anticipatory, outcome-driven account. This marks a major paradigm shift with respect to the cause of behaviour: away from stimulus-evoked, towards outcome-driven accounts of learned behaviour (Colwill & Rescorla 1990; Colwill & Motzkin 1994; Dickinson 2001; Elsner & Hommel 2001; Hoffmann 2003; McDannald *et al.* 2005). Does this paradigm shift need to be advocated for animals with a minimal (Ramaekers *et al.* 2005) nervous system as well? We tackle this issue in *Drosophila* larvae, usually regarded as simple 'feeding machines' with about 10 million times fewer neurons when compared to man.

In the wild, and occasionally in our kitchens, *Drosophila* larvae live in the superficial layers of rotting fruit. They are the feeding stages of the flies' life cycle and as such are largely concerned with feeding. In the laboratory, one can differentially condition them to associate the taste of a sweetened agarose substrate with one odour, and a non-sweetened agarose substrate with another odour. After such training, the larvae prefer the previously rewarded over the previously non-rewarded odour in a binary choice assay (Scherer *et al.* 2003; Hendel *et al.* 2005; Neuser *et al.* 2005; Gerber & Stocker in press). Canonically, such conditioned approach is explained by a changed value of the odour. After training, the odour can activate a set of modified synapses (similar to the situation in adult flies (Gerber *et al.* 2004), these probably are the output synapses of the larval mushroom bodies (Heisenberg *et al.* 1985; Honjo & Furukubo-Tokunaga 2005)) which provide input to premotor areas. After training, this odour-evoked input to the premotor areas is suggested to be sufficiently strong to trigger conditioned approach. In an obviously less parsimonious, outcome-driven account, one would have to suggest that conditioned approach is expressed because animals expect a benefit from tracking down the odour. Our experiments pit these accounts against each other.

2. MATERIAL AND METHODS

(a) *Conditioning*

For appetitive learning, 5-day-old feeding stage larvae receive either of two training regimes: amylacetate (AM) is presented with reward and 1-octanol (OCT) without reward (AM+/OCT), or they are trained reciprocally (AM/OCT+). For aversive learning, the procedure is analogous (AM−/OCT or AM/OCT−). In the test, we always measure the choice between AM and OCT.

Petri dishes (85 mm in diameter) are filled with 1% agarose and used the following day. As reward, we use 2 mol of fructose stirred with 1 l of agarose. As punishment, we use quinine hemisulphate (0.2%) or sodium chloride (4.0, 0.5 or 0.375 M).

Experiments are performed in red light under a fume hood. Perforated Teflon containers are loaded with 10 µl of odourant (AM diluted 1 : 50 in paraffin oil or OCT) and placed onto the assay plate, which may or may not contain a reinforcer. Thirty larvae are transferred to the assay plate, and after 5 min they are transferred to a fresh plate with the alternative odourant–substrate combination. This cycle is repeated three

* Author for correspondence (bertram.gerber@biozentrum.uni-wuerzburg.de).
† Present address: Max Planck Institute of Neurobiology, Department of Systems and Computational Neurobiology, Am Klopferspitz 18a, 82152 Martinsried, Germany

times. Then, animals are placed in the middle of an assay plate with AM on one side and OCT on the other. This test plate has no reinforcer added, unless noted otherwise.

(b) Measures and statistics

After 3 min, we calculate an odour preference, PREF $(-1; 1)$ as the number of animals at the AM side ($\#_{AM}$) minus the number of animals at the OCT side ($\#_{OCT}$), divided by the total ($\#_{TOTAL}$):

(i) $\text{PREF} = (\#_{AM} - \#_{OCT})/\#_{TOTAL}$

From alternately run, reciprocally trained groups we calculate a learning index (LI) $(-1; 1)$:

(ii-a) $\text{LI} = (\text{PREF}_{AM+/OCT} - \text{PREF}_{AM/OCT+})/2$
(ii-b) $\text{LI} = (\text{PREF}_{AM-/OCT} - \text{PREF}_{AM/OCT-})/2$

Thus, positive LIs indicate appetitive memory, whereas negative LIs indicate aversive memory (for a detailed discussion of these calculations see appendix in Hendel *et al.* (2005)). Non-parametric statistics (one-sample sign test, Kruskal–Wallis test and Mann–Whitney U-test) are used throughout (p-value 0.05) and results of these tests are shown in figures 1 and 2.

3. RESULTS

Outcome-driven models of behaviour control suggest that behaviours are expressed if their outcomes are desired (Hoffmann 2003). Consider that after differential conditioning of one odour with sugar and another without sugar, larvae find themselves in a binary choice situation with one odour suggesting 'here you will find sugar', whereas the alternative suggests 'here you will not find sugar'. In the absence of sugar, larvae should go towards that odour which suggests the desired outcome, i.e. towards the sugar-associated odour. If sugar is already present, tracking down that odour is pointless. In contrast, after aversive training with, for example, an unsavoury concentration of salt, one odour suggests 'here you will suffer from salt', whereas the alternative suggests 'here you will not suffer from salt'. Thus, if salt is absent, tracking down the no-salt associated odour is pointless. However, in the presence of salt, tracking down the no-salt associated odour can lead to the desired outcome, i.e. relief from salt; therefore, conditioned behaviour should be expressed. In short, this suggests that appetitive memories in larval *Drosophila* are behaviourally expressed only in the absence of the appetitive reinforcer (search behaviour for the predicted reward), whereas aversive memories should be expressed only in the presence of the aversive reinforcer (flight behaviour to escape the aversive reinforcer). This seems reasonable, as searching for something that is present, or escaping from something that is actually absent would seem eccentric at best.

We test these predictions by discriminatively training fruit fly larvae to associate an odour with either sugar, a bitter reinforcer or a salty reinforcer (the latter at either high, medium or low concentration; this classification is based on the relative preference between bitter and salt: for the high salt concentration, the larvae prefer bitter; for the low salt concentration, they prefer salt; see fig. S1 in the electronic supplementary material). A second odour is always presented without any reinforcer. We then test the
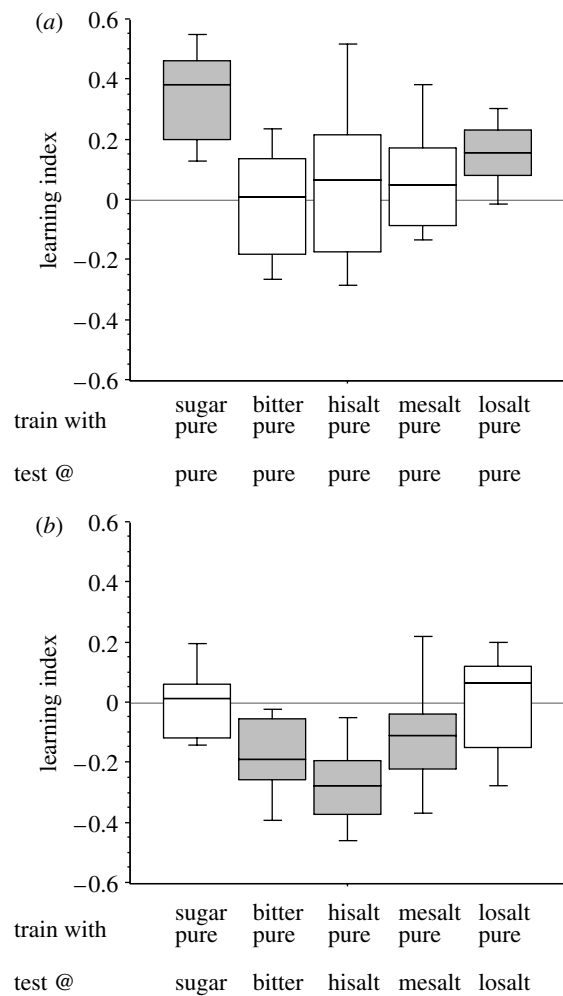
Figure 1. Memory expression in the absence or presence of the training-reinforcer. (*a*) In the absence of the training-reinforcer, appetitive memories are expressed, but aversive memories are not expressed. (*b*) In contrast, in the presence of the training-reinforcer, appetitive memories are not expressed, but aversive memories are expressed. Positive LI scores indicate appetitive memory, whereas negative LI scores indicate aversive memory. The box plots represent the median LI as the bold line, 25% and 75% quantiles as the box boundaries, and 10% and 90% quantiles as whiskers. In both (*a*) and (*b*), learning indices differ between groups (Kruskal–Wallis test; (*a*) $p < 0.05$, $H = 26.4$, d.f. $= 4$, N from left to right: 19, 16, 12, 16, 13; (*b*) $p < 0.05$, $H = 20.9$, d.f. $= 4$, N from left to right: 11, 16, 12, 18, 13). Significant differences from zero ($p < 0.05$, one-sample sign tests) are indicated by the shading of the boxes. U-tests for 'vertical' comparisons between the columns in (*a*) versus (*b*) are from left to right: $U = 9$, $N = 19$, 11; $U = 62$, $N = 16$, 16; $U = 24$, $N = 12$, 12; $U = 71$, $N = 16$, 18; $U = 43$, $N = 13$, 13; $p < 0.05$ in all cases. Hisalt, high salt; mesalt, medium salt; losalt, low salt.

choice between the two odours in either the absence (figure 1*a*) or presence (figure 1*b*) of that reinforcer which had been used for training. If the training-reinforcer is absent during test (figure 1*a*), the larvae behaviourally express appetitive memory after sugar as well as after low-salt training. However, after aversive training with either bitter, high-salt or medium-salt, animals do not behaviourally express any memory. These findings replicate those in Hendel *et al.* (2005). However, if the training-reinforcer is present during test, we find the inverted pattern of results (figure 1*b*), i.e. animals show no appetitive memory in the presence of the appetitive reinforcer, whereas they show

aversive memory in the presence of the aversive training reinforcer. Thus, in line with the concept of anticipatory, outcome-driven behaviour control, the animals seem to express their memory in behaviour only when doing so promises a positive outcome; more specifically, it seems critical that the behaviour in question can be expected to improve the current situation.

Is it possible to account for our data by assuming that the reinforcers act as retrieval cues at test? If they do, memory scores should be higher when the test situation is more similar to the training situation. Thus, in the presence of the training-reinforcer, memory scores should always be higher than in the absence of it. Our results concerning appetitive reinforcement using sugar and low salt (compare figure 1a,b) refute this notion.

What about direct effects of the reinforcers during training? Maybe the experience with the reinforcers during training determines memory scores solely by inducing a permissive state which carries over to the test or which determines the levels of learning (Pompilio et al. 2006)? If this were true, animals should show the same memory score after the same kind of training, irrespective of the test situation. This is refuted by the observation that in all 'vertical' comparisons in figure 1 (and also in figure 2), i.e. between pairs of groups that have undergone the same training, we find differences in memory scores.

Can one argue that the reinforcers have direct effects of another kind? Maybe the worse the situation during the test, the more strongly the animals express their memory in behaviour ('if in trouble, use your brain')? Thus, if the test situation is permissive, memory should be expressed. However, groups that are tested under equal conditions in figure 1a (and also in figure 2a) differ in terms of memory scores. This suggests that such a notion cannot account for our data. Thus, the concept of anticipatory, outcome-driven behaviour control can fully account for the present results, but assuming either a role of the reinforcers as retrieval cues or direct effects of the reinforcers, during training or test, cannot account for these results.

In the next experiment, we further scrutinize predictions from the concept of outcome-driven behaviour control. We trained three groups of larvae such that for all groups, one odour is presented with bitter and the other odour with salt; groups differ with respect to the concentration of salt, which was either high, medium or low. Then, all groups are tested in the presence of bitter. If the motivational state, as induced by the test situation (or the similarity between training and test situation; see above) were the sole determinant of memory expression, then all groups should express memory. This is not the case, as memory scores differ between these groups (figure 2a); only the groups trained with bitter/medium salt and bitter/low salt show significant aversive memory scores for the bitter-associated odour, whereas the group trained with bitter/high salt does not show such scores. These results are readily explained by suggesting that memories are behaviourally expressed only when doing so can improve the situation, i.e. if trained with bitter/high salt, bitter is the less bad of the two options (see Fig. S1 in the electronic supplementary material). Therefore, in the presence of bitter, no memories are expressed. As the salt concentration is reduced, bitter becomes the worse of the two options (see Fig. S1 in the electronic supplementary material), and hence the larvae start to show their
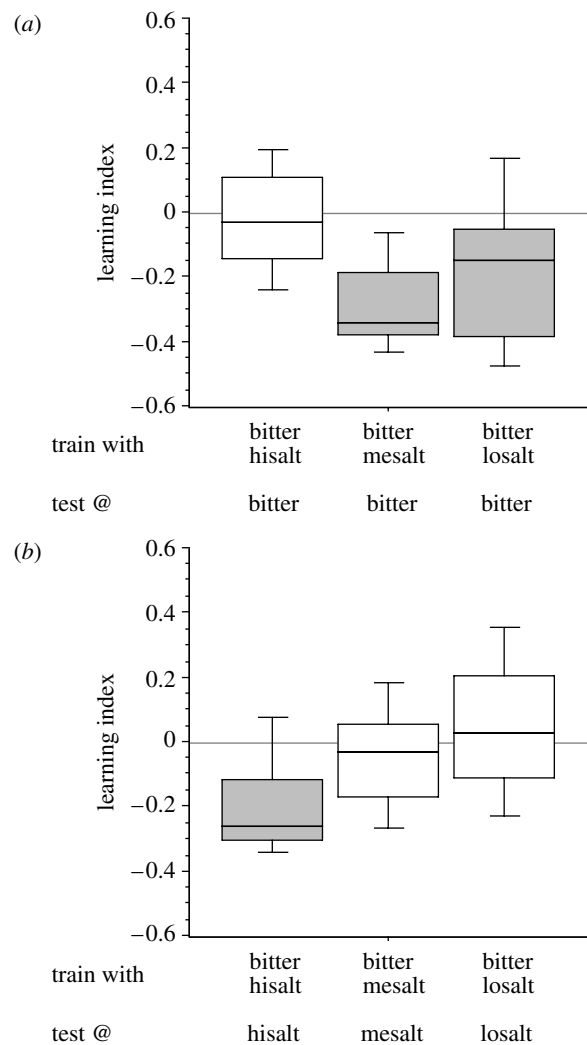

Figure 2. Memory expression in the presence of (a) bitter or (b) various concentrations of salt during test. All groups are trained such that one odour is paired with bitter and the other with salt; groups differ with respect to the concentration of salt which was either high, medium or low. Testing was then performed either in the presence of (a) bitter or (b) that salt concentration which had been used for training. Depending on the substrate on which the test was performed, the inverse pattern of behavioural memory expression is found. In both (a) and (b), learning indices differ between groups (Kruskal–Wallis test; (a) $p < 0.05$, $H = 8.0$, d.f. = 2; N from left to right: 12, 11, 22; (b) $p < 0.05$, $H = 11.2$, d.f. = 2; N from left to right: 12, 12, 22). Significant differences from zero ($p < 0.05$, one-sample sign tests) are indicated by the shading of the boxes. Results of U-tests for 'vertical' comparisons between the columns in (a) versus (b) are, from left to right: $U = 31$, $N = 12, 12$; $U = 20$, $N = 11, 12$; $U = 119$, $N = 22, 22$; $p < 0.05$ in all cases. Abbreviations as in figure 1.

memories in the presence of bitter. This predicts that the pattern of memory scores should be inverted if animals are tested in the presence of the respective salt concentration: if trained with bitter/high salt, high salt is the worse of the two options, and hence, in the presence of high-salt, memory should be expressed. As the salt concentration is reduced, salt becomes the better of the two options; therefore, in the presence of these lower salt concentrations, memories should not be observed. This is indeed what we find (figure 2b).

A final argument is derived by considering the medium salt concentration, which can induce aversive memories,

and which does permit expressing significant memory scores when present during the test (figure 1*b*). However, medium-salt memory is not expressed in figure 2*b*. This, we argue, is because medium salt in figure 1*b* is the worse of the two options, but not in figure 2*b*. Therefore, it is not the value of the test situation *per se* which determines the behavioural expression of memory.

Obviously, memory scores are determined neither by the strength of the established memory trace alone, nor by the value of the test situation alone, but by their interaction. We argue that this interaction between what 'may be' (based on olfactory memory) and what 'is' (based directly on gustatory input) can provide the animals with an estimate of their behaviour's expected outcome.

4. DISCUSSION

We therefore suggest that when presented with the choice between the previously reinforced and the previously non-reinforced odour at the moment of testing, memory usage in larval *Drosophila* involves a two-step process. In the first step, irrespective of the test situation, the odour activates its memory trace. In the second, hitherto unrecognized evaluative step, a comparison is made between the value of this activated olfactory memory trace and the value of the current situation. If the value of the odour memory is higher than that of the current situation, tracking down that odour can be expected to improve the situation; thus, memory will be expressed in terms of appetitive search for the predicted reward. If the current situation is equal to or better than what the odour memory suggests, then tracking down that odour cannot be expected to lead to any improvement and no memory will be observable in behaviour. In contrast, aversive memories lead to a conditioned flight response only if the test situation requires flight. In other words, the 'expected outcome' is computed as the difference between two pieces of readily available information: the value of the activated memory trace and the value of the current situation. It is this expected outcome, rather than the activated memory trace *per se*, which is the immediate cause of conditioned behaviour.

The recent discovery that stimulation of octopaminergic/tyraminergic neurons in the present larval learning paradigm can substitute for appetitive reinforcement during training and that stimulation of dopaminergic neurons can in turn substitute for aversive reinforcement (Schroll *et al*. 2006) raises the question whether these neuromodulators may also be involved in the computation of expected outcomes during test.

REFERENCES

Colwill, R. M. & Motzkin, D. K. 1994 Encoding of the unconditioned stimulus in Pavlovian conditioning. *Anim. Learn. Behav.* **22**, 384–394.

Colwill, R. M. & Rescorla, R. A. 1990 Effect of reinforcer devaluation on discriminative control of instrumental behavior. *J. Exp. Psychol. Anim. Behav. Process.* **16**, 40–47. (doi:10.1037/0097-7403.16.1.40)

Dickinson, A. 2001 The 28th Bartlett Memorial Lecture: causal learning: an associative analysis. *Q. J. Exp. Psychol. B* **54**, 3–25. (doi:10.1080/02724990042000010)

Elsner, B. & Hommel, B. 2001 Effect anticipation and action control. *J. Exp. Psychol. Hum. Percept. Perform.* **27**, 229–240. (doi:10.1037/0096-1523.27.1.229)

Gerber, B. & Stocker, R. F. In press. The *Drosophila* larva as a model for studying chemosensation and chemosensory learning: a review. *Chem. Sens.*

Gerber, B., Tanimoto, H. & Heisenberg, M. 2004 An engram found? Evaluating the evidence from fruit flies. *Curr. Opin. Neurobiol.* **14**, 737–744. (doi:10.1016/j.conb.2004.10.014)

Heisenberg, M., Borst, A., Wagner, S. & Byers, D. 1985 *Drosophila* mushroom body mutants are deficient in olfctory learning. *J. Neurogen.* **2**, 1–30.

Hendel, T. *et al*. 2005 The carrot, not the stick: appetitive rather than aversive gustatory stimuli support associative olfactory learning in individually assayed *Drosophila* larvae. *J. Comp. Physiol. A.* **191**, 265–279. (doi:10.1007/s00359-004-0574-8)

Hoffmann, J. 2003 Anticipatory behavioral control. In *Anticipatory behavior in adaptive learning systems* (ed. M. V. Butz, O. Sigaud & P. Gerad), pp. 44–65. Heidelberg, NY: Springer.

Honjo, K. & Furukubo-Tokunaga, K. 2005 Induction of cAMP response element-binding protein-dependent medium-term memory by appetitive gustatory reinforcement in *Drosophila* larvae. *J. Neurosci.* **25**, 7905–7913. (doi:10.1523/JNEUROSCI.2135-05.2005)

McDannald, M. A., Saddoris, M. P., Gallagher, M. & Holland, P. C. 2005 Lesions of orbitofrontal cortex impair rats' differential outcome expectancy learning but not conditioned stimulus-potentiated feeding. *J. Neurosci.* **25**, 4626–4632. (doi:10.1523/JNEUROSCI.5301-04.2005)

Neuser, K., Husse, J., Stock, P. & Gerber, B. 2005 Appetitive olfactory learning in *Drosophila* larvae: effects of repetition, reward strength, age, gender, assay type and memory span. *Anim. Behav.* **69**, 891–898. (doi:10.1016/j.anbehav.2004.06.013)

Pompilio, L., Kacelnik, A. & Behmer, S. T. 2006 State-dependent learned valuation drives choice in an invertebrate. *Science* **311**, 1613–1615. (doi:10.1126/science.1123924)

Ramaekers, A., Magnenat, E., Marin, E. C., Gendre, N., Jefferis, G. S., Luo, L. & Stocker, R. F. 2005 Glomerular maps without cellular redundancy at successive levels of the *Drosophila* larval olfactory circuit. *Curr. Biol.* **15**, 982–992. (doi:10.1016/j.cub.2005.04.032)

Rescorla, R. A. 1988 Behavioral studies of Pavlovian conditioning. *Annu. Rev. Neurosci.* **11**, 329–352. (doi:10.1146/annurev.ne.11.030188.001553)

Rescorla, R. A. & Wagner, A. R. 1972 A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In *Classical conditioning II: current research and theory* (ed. A. H. Black & W. F. Prokasy), pp. 64–99. New York, NY: Appleton-Century-Crofts.

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

Schroll, C. *et al*. 2006 Light-induced activation of distinct modulatory neurons substitutes for appetitive or aversive reinforcement during associative learning in larval *Drosophila*. *Cur. Biol.* (in press).