Reward learning in normal and mutant Drosophila
(genetics/memory/positive reinforcement)

BRUCE L. TEMPLE*, NANCY BONINI†, DOUGLAS R. DAWSON‡, AND WILLIAM G. QUINN

Department of Biology, Princeton University, Princeton, New Jersey 08544

Communicated by David H. Hubel, December 1, 1982

ABSTRACT Hungry fruit flies can be trained by exposing them to two chemical odorants, one paired with the opportunity to feed on 1 M sucrose. On later testing, when given a choice between odorants the flies migrate specifically toward the sucrose-paired odor. This appetitively reinforced learning by the flies is similar in strength and character to previously demonstrated negatively reinforced learning, but it differs in several properties. Both memory consolidation and memory decay proceed relatively slowly after training with sucrose reward. Consolidation of learned information into anesthesia-resistant long-term memory requires about 100 min after training with sucrose compared to about 30 min after training with electric shock. Memory in wild-type flies persists for 24 hr after training with sucrose compared to 4–6 hr after training with electric shock. Memory in amnesiac mutants appears to be similarly lengthened, from 1 hr to 6 hr, by substituting sucrose reward for shock punishment. Two other mutants, dunce and rutabaga, which were isolated because they failed to learn the shock-avoidance task, learn normally in response to sucrose reward but forget rapidly afterward. One mutant, turnip, does not learn in either paradigm. Reward and punishment can be combined in olfactory discrimination training by pairing one odor to sucrose and the other to electric shock. In this situation, the expression of learning is approximately the sum of that obtained by using either reinforcement alone. After such training, memory decays at two distinct rates, each characteristic of one type of reinforcement.

We hope to further our understanding of learning and memory by analyzing genetic mutations that disrupt these processes. Fruit flies can learn various tasks after conditioning with negative reinforcement (1–4). For example, they can discriminate between two odors and specifically avoid one that they have previously experienced in association with electric shock (1). Several single-gene Drosophila mutants either fail to learn this negatively reinforced task or forget it rapidly (3, 5, 6).

Here we report a similar learning situation involving positive reinforcement, in which hungry flies learn to run toward specific odors that they have experienced in association with the opportunity to feed. We examine this appetitive learning and the subsequent memory, compare it with negatively reinforced learning, and test the behavior of mutants in this new learning situation.

MATERIALS AND METHODS

Stocks and Conditions. The wild-type strain used here (and the parent strain for all mutants) was Canton-S (C-S). Culture methods and experimental equipment were as described (1, 5). Before training, populations of flies were starved for 18–20 hr in pint bottles with two 5-cm Whatman no. 1 filter paper discs and 1 ml of H₂O. This leaves them hungry but apparently healthy.

Appetitive Learning. Fifteen minutes before training, a solution of odorant—either 0.5% 3-octanol (OCT) or 1% 4-methylocyclohexanol (MCH)—was spread on printed circuit grids (Fig. 1A). Grids to provide positive reinforcement had, in addition, a 1-cm-wide band of 1.0 M sucrose solution halfway along their length, applied after the odorant. Tubes with appropriate grids were arranged in the training apparatus as shown in Fig. 1B. About 40 flies were introduced into the rest tube and given 1 min to become acclimated; then they were tapped down into a small cylindrical “start chamber” (1 cm in diameter, 1 cm deep) in which they could be transferred into alignment with the appropriate grid tubes. During training, the apparatus was upright with three 18-inch (46-cm) cool white fluorescent light sources giving 1.1 × 10⁻⁴ W/cm² from above and 1.5 × 10⁻⁴ W/cm² from the side opposite the experimenter. Driven by phototaxis and negative geotaxis, the flies ran from the start chamber up into the grid tubes. A training cycle consisted of sequential 30-sec exposures to odor A (no reward), rest, odor B (reward), and rest. This cycle was run twice.

The flies’ odor preference was tested by transferring them to a choice chamber (Fig. 1C) that was rigidly positioned relative to the light sources. One minute after training, they were tapped into the sliding center compartment and transported to a point between grid tubes containing odors OCT and MCH, both presented without reinforcement. The flies were given 15 sec to distribute themselves and then the center compartment was slid back, holding them in the tubes they had chosen. Flies in each tube and in the center compartment were then counted. In the second half of the experiment, a new group of flies was trained as above except that sucrose was paired with the other odor. These flies were tested in the same choice chamber apparatus used to test the first group. The index of learning we use for discriminative training, A, is the fraction of flies choosing the sucrose-associated odor minus the fraction choosing the control odor, averaged for reciprocal halves of the experiment (1). A values were calculated for each of 8–12 complete experiments run as above and then averaged to give the mean ± SEM reported here.

The arrangement of training and testing tubes in the apparatuses was actually systematically permuted to avoid giving inadvertent procedural cues to the flies. The reward-learning effect and the long memory were demonstrated in blind procedures; the data are included here. Experiments comparing mutant and wild-type learning were performed with the experimenter ignorant of the genotype being tested.

Shock-Avoidance Learning. Training to electric shock and subsequent testing were done as described (1) except that we

Abbreviations: C-S, Canton-S wild-type fly strain; OCT, 3-octanol; MCH, 4-methylocyclohexanol; A, index of learning.

* Present address: Dept. of Biochemistry, George Washington Univ. Medical School, Washington, DC 20037.
† Present address: Dept. of Biochemistry, Univ. of Wisconsin, Madison, WI 53706.
‡ Present address: Dept. of Psychology, Univ. of California, Irvine, CA 92717.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.
used two training cycles rather than three, to make the avoidance training procedure more similar to the appetitive training procedure above.

Memory. To test memory periods of 12 hr or less, flies were starved for 19 hr, trained as usual, and then kept in test tubes with foam plugs in a humidified chamber until testing time. For longer memory periods, the flies were given standard cornmeal food for 1–6 hr after training. They were then transferred to test tubes in a humidifier and starved for 17–19 hr before testing.

To measure consolidation of learned information into long-term memory—i.e., memory that is relatively resistant to anesthesia—we subjected groups of trained flies in test tubes to brief cooling (1 min at 4°C) at various times after training. This produced rapid and reversible anesthesia. The methods of cooling were essentially as described (7). However, we obtained better results after two-cycle shock-avoidance training if we cooled the flies in the plastic tubes used for training, omitting transfers to and from glass tubes. Therefore, cold anesthetization after shock-avoidance training was done in 17 × 100 mm plastic tubes. This variant anesthetization procedure did not affect results after sucrose reward training, and about 25% of these experiments used plastic tubes.

Statistical Analysis. In two-odor discriminative-learning runs, the significances of associative learning effects (positive As) were assessed with one-tailed t tests. Differences in As between normal and mutant flies, or between flies after different training regimens were assessed with two-tailed t tests. In one-odor training, the significance of the associative effect relative to all controls was measured by using a one-way analysis of variance with a Neuman–Keuls test. In Results, all statements about learning or differences in learning are significant (P < 0.01) unless otherwise indicated.

RESULTS

Learning with Reward. Hungry Drosophila can learn to associate specific odors with food reward. When the flies were tested in a choice-chamber T maze, 1 min after training, most of them ran to the sucrose-associated odorant. When sucrose was paired with OCT during training, 57 ± 4% of the flies ran to the arm with OCT during testing and 33 ± 5% ran to the MCH arm; A = 0.24 ± 0.09. On the other hand, when MCH was paired with sucrose on training, on testing 70 ± 4% of the flies ran to MCH and 21 ± 4% of the flies ran to OCT; A = 0.49 ± 0.08. The mean A for the complete experiments was 0.36 ± 0.04.

C-S flies learned poorly without starvation (A = 0.09 ± 0.05). They learned optimally after 19–20 hr of starvation (A = 0.36 ± 0.04). Such starvation affected neither their intrinsic odor preference nor their learning scores in shock-avoidance training (data not shown). Longer starvation impaired learning, however, even in apparently healthy flies.

We ran a number of control experiments to be sure the odor preference we observed was genuine learning that was mediated by reward. First, note that our basic experimental design for discriminative learning was symmetrical, with two groups of flies trained to choose opposite odors in different halves of an experiment and tested in the same apparatus. Only the temporal pairing of odor with reinforcement was changed between the two halves of an experiment, and A, the learning index, measured only the difference in the flies’ behavior between the two halves. This should eliminate simple odor bias or nonassociative effects such as sensitization as explanations for the learning effect. Nevertheless, we worried that the band of sucrose in one of the training tubes might distort this symmetry by decreasing the ambient odor concentration in that tube and thus indirectly affecting the flies’ subsequent behavior. To rule out this possibility we replaced the sucrose solution with 1.0 M D-sorbitol, a sugar that is tasteless to the flies (8), and trained the flies as usual. When tested afterward, these flies showed no learning (A = −0.03 ± 0.05, n = 4). The sorbitol should block as much odorant as sucrose. Thus, the slight residual asymmetry in our experimental design cannot account for the learning effect normally observed.

Reward itself in some cases may cause a change in the aver-siveness of one or the other odor regardless of the temporal association between the reward and the odor (9). In our experimental conditions it did not: flies exposed to sucrose alone showed the same slight intrinsic preference for OCT over MCH (60 ± 4%) as totally naive flies (62 ± 3%). Moreover, when two-odor training was carried out as usual but with sucrose reward explicitly unpaired—i.e., separated from exposure to both odorants by 30-sec rest periods—the learning effect disappeared (A = −0.01 ± 0.07).

A genuinely learned response should disappear when the cues are presented without reinforcement. After discriminative con-
ditioning to an odor as usual, a single 30-sec exposure to that odor without reinforcement extinguished the learned response (Λ = 0.03 ± 0.04). It is worth noting that this is much more rapid than with negatively reinforced learning for which complete extinction requires 7–10 unreinforced odor exposures (1, 10).

Ordinarily we trained and tested about 40 flies at a time. This allowed rapid measurement of behavioral effects but could leave one uncertain about what individual flies were doing. Therefore we thought it was important to make sure that the basic learning effect was present when flies were tested individually. Small groups (five to eight) of flies were trained as usual to either OCT or MCH, then the flies were tested one by one in a choice between OCT and MCH. In tabulating this result, to obtain an average A with SEM we divided the 125 individuals into 12 groups, each composed of about 10 individuals that had been tested in the same choice chamber apparatus. A for individually tested flies was 0.37 ± 0.06, similar to that of flies tested en masse (0.36 ± 0.04). This similarity of behavior suggests that the flies in our groups behaved independently. There is no evidence for "guide flies."

We wanted to be certain that sucrose was acting as a genuine reward. Therefore, we ran some experiments in which flies were exposed to a single odor, alone, or in association with sucrose solution, and were given a choice between a tube with that odor and a neutral tube ("air") with no odorant added. Both OCT and MCH were aversive to naive flies, and OCT became increasingly aversive with repeated exposures (see also ref. 10). Prior exposure to sucrose did not change the flies' usual aversion to the odors on subsequent testing. However, when flies received two exposures to MCH paired with sucrose and fed on the sucrose both times, on later testing 65 ± 4% chose the MCH over air. Similarly, when OCT was paired with sucrose, 59 ± 3% chose OCT over air (Fig. 2). Both numbers are significantly greater than 50% (P < 0.01). Sucrose evidently does act as a reward, making a previously aversive odor attractive to the flies by association.

Memory. In our learning tests, the consequences of positive and negative reinforcement in altering the attractiveness of an odor to flies were roughly similar in magnitude, although opposite in sign. Pairing the odorant with sucrose made it more attractive in a subsequent test; pairing it with a 90-V electric shock, under somewhat different conditions, made it less attractive by about the same amount. Expressed in terms of As, the value after sucrose reward training (0.36 ± 0.04) was similar to that after shock-avoidance training (0.36 ± 0.03).

This rough equivalence held true only for learning effects measured immediately after training. In our experiments, positively reinforced memories persisted much longer than negatively reinforced ones. After two cycles of training, shock-reinforced memory decayed in 4–6 hr (Fig. 3) (10). Memory after sucrose-reinforced training persisted much longer and was appreciable at 24 hr (Λ = 0.15 ± 0.06).

The disparity in memory spans produced by the two reinforcements is dramatic. It suggested to us that the character of the reinforcement used might affect the way memory is stored. This issue is difficult to decide; there are certainly alternative explanations for the different memory spans. Of these, the most likely is that feeding on sucrose simply provided quantitatively stronger reinforcement, so that associated odors imprinted themselves more firmly on the fly's memory. If this is the explanation, one should be able to weaken the memory, and consequently shorten the memory span, by diluting the sucrose solution used in training. When we measured both immediate learning and memory 1–6 hr later as a function of sucrose concentration used in training, we found that dilution of the sucrose had virtually no effect on rate of memory decay, although at relatively dilute sucrose concentrations (0.025 M) there was some effect on initial learning scores (Fig. 4). We observed analogous results with negatively reinforced training. Here we varied the shock voltage between 10 V and 90 V and found that this had little effect on memory decay, although low voltages elicited less initial learning (Fig. 4). The difference in memory retention in sucrose-associated and shock-associated memories is not simply due to quantitative differences in the strength of the reinforcement provided. From the evidence at hand, there may be a qualitative difference in the way such associations are processed or stored.

Memory Consolidation. Immediately after training, memory in various animals is susceptible to various agents such as electroconvulsive shock and anesthetics which interfere with neural

---

**Fig. 2.** Learning after training to a single odor. Bars: A, naive; B, odor alone; C, sucrose alone; D, sucrose and odor, unpaired; E, sucrose and odor, paired. Flies experiencing either OCT alone or MCH alone avoided the odor when given a T-maze choice between a tube with that odorant and an unscented "air" tube. Flies exposed to sucrose alone or to sucrose explicitly unpaired by 30 sec to either MCH (black bars) or OCT (white bars) also avoided the odor being tested. Only when sucrose was paired with an odor was migration toward that odor significantly enhanced (P < 0.01 for OCT; P < 0.05 for MCH). In both these cases, >50% of the flies chose the reinforced odor over "air" (P < 0.01).

**Fig. 3.** Memory retention by normal C-S flies at various times after positively reinforced conditioning (●), negatively reinforced conditioning (○), and conditioning during which both reinforcements were presented (λ). Error bars show SEM for 6–12 experiments.
activity. Typically, such memory becomes consolidated later into a long-term form which is resistant to these agents. Our flies are conveniently anesthetized by cooling them, and earlier experiments (7, 10) have shown that they, like mammals, have both short- and long-term memory after shock-avoidance training with a characteristic time course for consolidation. We repeated these experiments under our present training conditions, anesthetizing flies at various times after training. We found that the shock-reinforced association became resistant to the treatment about 30–40 min after training whereas consolidation of sucrose-reinforced associations occurred 90–120 min after training (Fig. 5). Memory consolidation, like memory decay, proceeded more slowly after training with reward.

**Dual-Reinforcement Training.** Training with sucrose and with electric shock apparently produced learned associations with different physiological characteristics. We wondered how they would interact. Therefore, we combined positive and negative reinforcement in the same training session by electrifying (90 V) the grid with odor A, putting sucrose on the grid with odor B, and lengthening the rest periods between odor exposures to 60 sec. The effects of the two reinforcements were roughly additive—i.e., the learning elicited by both together (Δ = 0.62 ± 0.04) was nearly the sum (Δ = 0.72 ± 0.05) of each reinforcement used alone (Fig. 3).

**Mutants.** When we had established and characterized appetitive learning in normal flies, we used the procedure to test several *Drosophila* mutants selected for inability to learn or remember in a negatively reinforced task. The experimental picture is made clearest by considering *annesiac*. This mutant has normal learning after shock-avoidance training, but afterward it forgets about 4 times as fast as normal flies and its memory is gone within 1 hr (6). After training with sucrose, memory in *annesiac* persisted much longer (6–9 hr), although not as long as in normal flies after sucrose training (Fig. 6). It is as if training with reward stretches out subsequent memory proportionally in normal and *annesiac* flies.

Two other mutants, *dunce* and *rutabaga*, showed practically no learning with aversive training (Δ = 0.04 ± 0.02 and Δ = 0.00 ± 0.02, respectively). When trained with sucrose reinforcement, they learned quite well. In fact, *dunce* showed normal learning. Subsequently, however, both mutants forgot rapidly, at least 25 times as fast as normal flies (Fig. 6). Why do these mutants show learning deficits in one training situation and memory deficits in another? Our best guess is that they do learn after electric-shock training but their memory is so transient or so labile as to be undetectable on testing 60 sec later. In fact, Dudai (11, 12) reported evidence for such a fleeting memory in both these mutants.

Another learning mutant, *turnip*, showed undetectable learning with both the aversive (Δ = 0.02 ± 0.04) and appetitive learning (Δ = 0.01 ± 0.05). Mutations in the dopa decarboxylase gene also block both types of learning, affecting learning acquisition more than memory retention (13). These dopa decarboxylase mutations affect synthesis of two neurotransmitters,
dopamine and serotonin (14). In summary, mutations that blocked learning in the original negatively reinforced paradigm also affected performance in the new positively reinforced task. However, some of the mutations affected retention rather than immediate learning.

DISCUSSION

Fruit flies change their response to an odor after experiencing it in association with sucrose solution. This behavioral change results from associative learning; control experiments eliminate habituation, sensitization, odor bias, or group effects as explanations for the change. During training, the sucrose acts as a genuine reward, mediating appetitive learning. We infer this because the flies must be hungry for the behavioral change to take place. Moreover, after association with sucrose, previously aversive odorants become attractive to the flies, which migrate specifically toward them and prefer them to clean air.

Various insects (especially bees) have been conditioned to reward (15–17), and some earlier experiments have been done on Drosophila. For example, flies’ ability to locate a food source in a T-maze improved with experience (18). In another test situation, after experience, negatively geotactic flies tended to choose the textured arm of a T-maze that allowed them to continue migrating upward (19). Unfortunately, the observed effects were weak in these cases, and the cues and reinforcements were not always well-defined. In our study of positively reinforced learning, we circumvented some of these problems by running separate controls and by training and testing in different apparatuses. Our salient findings, however, are with memory and with mutants.

After we train flies to a 1.0 M sucrose reward, their memory persists for at least 24 hr, much longer than shock reinforced memory. This difference in decay rates is not due simply to quantitative differences in strength of reinforcement because with both sucrose and electric shock the duration of memory was independent of stimulus strength (sugar concentration or shock voltage) over a broad range. Consolidation into long-term memory also occurs more slowly if sucrose is substituted for electric shock. Perhaps a number of processes, including memory consolidation and memory decay, proceed more slowly after training with sucrose. This idea is supported by the finding that the normally brief memory spans of the mutants dunc, rutabaga, and amnesiac are concomitantly lengthened after training with sucrose.

Flies can be trained to an olfactory discrimination task in which one odor is associated with sucrose reward and the other with electric shock punishment. After such “stick-and-carrot” training, the learning shown by the flies is approximately the sum of the effects produced by each reinforcement alone. Memory after stick-and-carrot training wanes in two distinct components, decaying for half an hour at a rapid rate characteristic of shock-reinforced memory and then abruptly changing to a slower decay rate characteristic of sucrose-reinforced memory. It is as if positively and negatively reinforced memories can be acquired additively and lost independently. This fact, taken together with the disparate properties of the two types of memory (susceptibilities to extinction; rates of consolidation and decay) suggests to us that positively reinforced and negatively reinforced associations are processed differently in the fly brain and are stored in ways that do not interfere with one another. The evidence for such separate memory “bins” is behavioral and indirect, but this represents the simplest explanation for all our findings.

Positively reinforced training gives new information about dunc and rutabaga. These mutations were isolated because they blocked learning in the original olfactory discrimination test with shock reinforcement (3, 5). Both mutants were later shown to exhibit a transient and labile form of memory, provided that some internal controls were left out of the testing procedure (11, 12). dunc and rutabaga flies performed with variable success on other negatively reinforced learning tests (4, 20, 21). Here we find that both mutants learn well after sucrose-reinforced training but show rapid memory decay afterward. The finding that the dunc and rutabaga mutations can specifically affect memory retention is particularly interesting because the biochemical deficit caused by each mutation is known. dunc flies lack one isozyme form of the enzyme cyclic AMP phosphodiesterase (22). rutabaga flies show a reduction in adenylyl cyclase enzyme activity (23). Results from Drosophila and also from Aplysia (24) suggest a strong link between memory and cyclic AMP metabolism.

We thank Seymour Benzer, Margaret Livingstone, and Christie Sahley for ideas and helpful criticism. We thank Joan Nielsen for typing many drafts. B.L.T. is a National Science Foundation Predoctoral Fellow. This work was supported by National Institutes of Health Grant GM 25578.