The biology of psychology

'Simple' conditioning?

Julien Colomb and Björn Brembs*

Freie Universität Berlin; Institut für Biologie-Neurobiologie; Berlin, Germany

Operant (instrumental) and classical (Pavlovian) conditioning are taught as the simplest forms of associative learning. Recent research in several invertebrate model systems has now accumulated evidence that the dichotomy is not as simple as it seemed. During operant learning in the fruit fly Drosophila, at least two genetically distinct learning systems interact dynamically. Inspired by analogous results in three other research fields, we propose to term one of these systems world-learning (assigning value to sensory stimuli) and the other self-learning (assigning value to a specific action or movement). During the goal-directed phase of operant learning, world-learning inhibits self-learning (in Drosophila via the mushroom-body neuropil), to allow for flexible generalization. Extended training overcomes this inhibition in a phase transition akin to habit formation in vertebrates, allowing self-learning to transform spontaneous actions to habitual responses. In part, these insights were achieved by reducing operant experiments beyond the traditional set-ups (i.e., 'pure' operant learning) and using modern, molecular and/or genetic model systems.

Differentiating Operant and Classical Conditioning

Every high-school student today learns about the dichotomy of simple conditioning experiments: Pavlov's dogs (classical or Pavlovian conditioning) and Skinner's rats (operant or instrumental conditioning). Classical and operant conditioning were recognized as producing two separate types of learning more than 70 years ago. Despite their clear procedural differences, it was recognized early on that the psychological processes occurring during conditioning were not as easily separable.[1-5] Specifically, classical associations between sensory stimuli were often found to be present after operant conditioning.[6] Conversely, as classical training progressed, operant processes were initially hypothesized to also occur as responding to the conditioned stimulus was rewarded by presentation of the unconditioned stimulus.[7] Already in these early days, it became obvious that the operational terms 'operant' and 'classical', while unambiguously distinguishing the two types of experiments, did not clearly delineate which processes actually occurred during learning. However, an experimental approach dissecting individual learning processes was not available and the debate lingered on.[8] Today, we can propose a terminology to better distinguish what is learned (stimuli or behavior) from how it is learned (by classical or operant conditioning). We define self-learning as the process of assigning value to a specific action or movement. We define world-learning as the process assigning value to sensory stimuli. While only world-learning occurs in classical conditioning experiments, both processes may occur during operant conditioning.

More than Synaptic Plasticity

In the torque meter apparatus,[9] Drosophila fruit flies can be subjected to different operant conditioning experiments (**Fig. 1**).[10] Using one of these operant paradigms to induce only world-learning (**Fig. 1A**;

there is no contingent relation between any specific behavior and punishment) and another operant paradigm to induce only self learning (**Fig. 1B**; no contingent stimuli present), we found that the cAMP pathway was necessary for world-learning, but dispensable for self-learning.[11] These results corroborate the evidence of the cAMP pathway being central to classical conditioning,[12] during which only world-learning is induced. In our setup, there is no residual performance in cAMP mutants, suggesting that this is the only pathway involved.[13] This contrasts with reports of olfactory classical conditioning, where cAMP-independent learning can be found. Most interestingly, inhibition of PKC activity affects self-learning, but not world-learning. Recent studies in Aplysia also imply a role of PKC in self-learning, suggesting that this separation may be evolutionarily conserved.[14] However, data involving PKC in mammalian world-learning indicate a dissociation on the level of the various PKC isoforms.[15,16] This double dissociation of cAMP and PKC in Drosophila has allowed us to use the two mechanisms as markers for world- and self-learning, enabling us to dissect the interaction between the two learning systems during operant conditioning situations in which both learning systems may be engaged (**Fig. 1C**). The vast majority of ethologically relevant learning situations can be classified as such composite situations.

Dynamic Learning Rates: Reciprocal Interactions between Multiple Learning Systems

For this dissection, we trained the animals in a composite operant task and then tested for evidence of world- and self-learning (**Fig. 2**). After the same amount of training sufficient to induce either world- or self-learning separately, flies only show evidence for world-learning.[17] This result suggests that during composite operant training, world-learning is preferentially engaged, while self-learning is suppressed. Interestingly, mutants in the cAMP pathway (i.e., with impaired world-learning), show no such suppression, suggesting that world-learning needs to be intact for the inhibiting effect to occur (**Fig. 2**). In wild

type animals, the inhibition appears to be overcome by intensive training (twice as long), because after such prolonged training flies do show evidence for self-learning (**Fig. 2**). This is reminiscent of studies in mammals, where the behavioral strategies used during test also depend on the amount of training. For instance, in navigation studies, relatively short training preferentially engages an allocentric strategy (the animal orients primarily according to environmental cues), while longer training induced an egocentric strategy (the animals performed the same sequence of movements).[18] The analogy to world- and self-learning is striking. The terminology of world- and self-learning itself was inspired by analogous developments in another research field.[19] There is a third field in which analogous results have been obtained. In experiments with rodents in operant chambers, extended training abolishes sensitivity to reinforcer devaluation by the process of habit formation which transforms goal-directed actions to habitual responses.[20] Also in this case, one can explain habit formation by the same interaction between world- and self-learning we discovered in flies. Habitual or compulsive behaviors can thus be considered as a particularly stable consequence of self-learning (**Fig. 3**). From this perspective, one may posit that habit formation requires repetition because it is inhibited by world-learning. After prolonged training, this inhibition is overcome and self-learning kicks in to form habits. We have discovered that in flies, a prominent neuropil, which is dispensable for both world- and self-learning, is involved in the inhibition of self-learning: the mushroom-bodies.[17]

A New Field of Research

These recent developments open a whole new field of research. A 'boxology' is often helpful to conceptualize the current working model (**Fig. 3**) and to guide these research efforts. Specifically, the biological basis of self-learning and its inhibition is still unexplored. A first step would be to identify the PKC isoform(s) required for self-learning and its targets in the neuron. While there is evidence from Aplysia that this PKC-dependent self-leaning

may involve neuron-wide plasticity,[21] it is not yet known if this is also the case in other organisms. Most interestingly, the mechanism of inhibition of self-learning by world-learning is yet to be investigated. It is tempting to speculate about a direct action of the cAMP pathway on the PKC pathway, supposing that both take place inside the same neurons. However, the mushroom-bodies are neither required for world- nor for self-learning but for the inhibitory interaction between the two. This implies that the inhibition depends on circuits distinct from the neurons where cAMP and PKC are acting. Once discovered, the mechanism by which extended training can overcome the inhibition may even be clinically relevant. This interaction seems to be a key step in the formation of habits and compulsive behaviors. Unraveling its mechanism may therefore help treating patients suffering from addiction or other compulsive disorders.

Last but not least, another major puzzle is a third process taking place in operant situations, which is involved in finding out which behavior controls which environmental stimuli, i.e., 'operant behavior'.[22] It is this little-understood process which is believed to underlie the generation-effect ("learning-by-doing"), i.e., the facilitation of world-learning by being in control of the stimuli which are to be learned.[23-28] Like the processes above, the mechanism by which this facilitation occurs remains elusive. The only results so far are negative: none of the mutants and transgenes tested in the last two decades shows any deficit in operant behavior. Nevertheless, modern molecular and genetic methods making it possible to study the function of lethal mutations in adult animals, are good reasons to be optimistic for the research on the mechanisms of operant behavior, despite these negative results.

References

1. Skinner BF. Two types of conditioned reflex and a pseudo type. Journal of General Psychology 1935; 12:66-77.
2. Konorski J, Miller S. On two types of conditioned reflex. Journal of General Psychology 1937; 16:264-72.
3. Skinner BF. Two types of conditioned reflex: A reply to Konorski and Miller. Journal of General Psychology 1937; 16:272-9.
4. Konorski J, Miller S. Further remarks on two types of conditioned reflex. Journal of General Psychology 1937; 17:405-7.
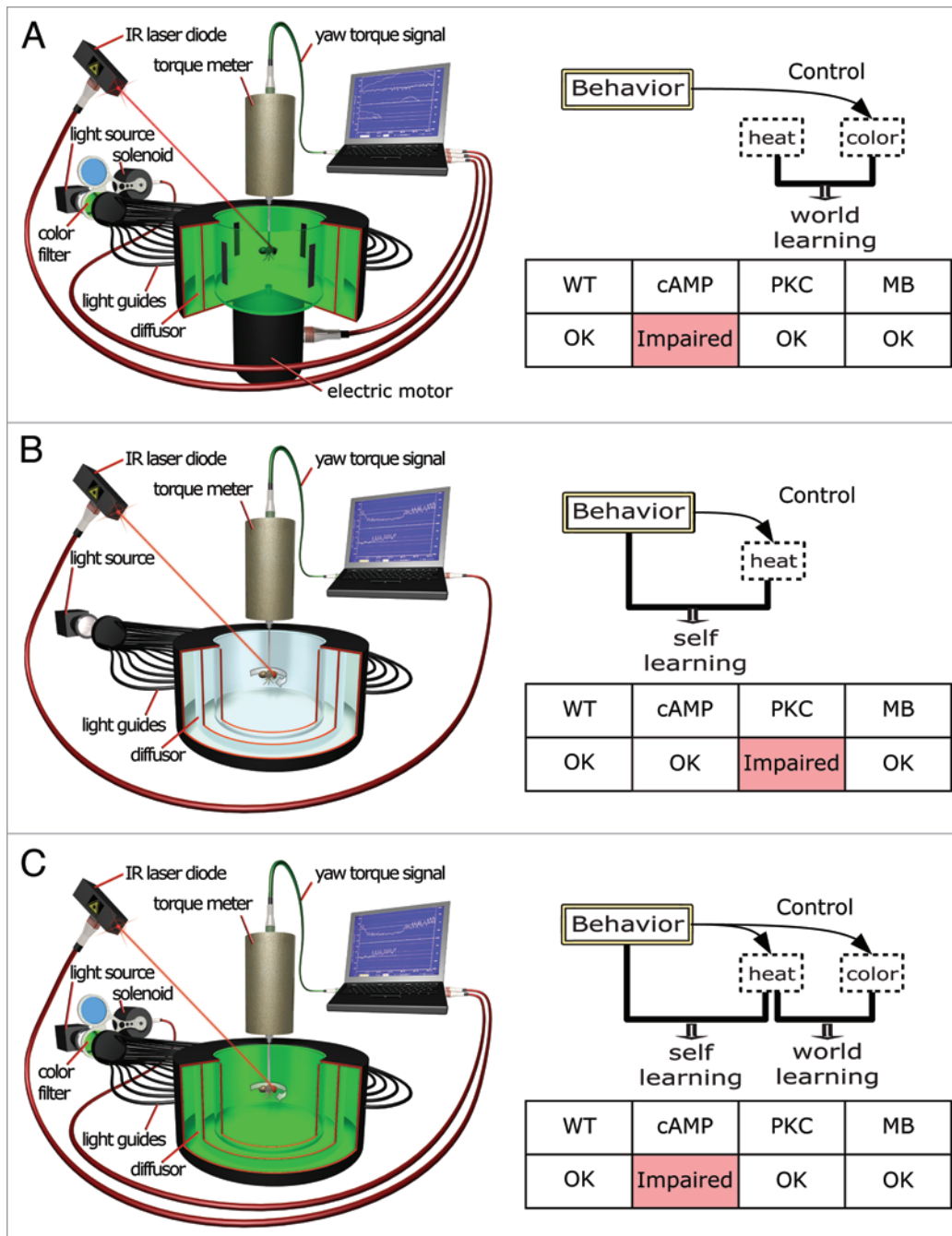
Figure 1. Three different operant conditioning procedures requiring different biological processes. Left - schematic representation of the experimental set-up: In all experiments, the flies are tethered to a torque meter which measures the angular momentum around the fly's vertical body axis (yaw torque), caused by attempted turning maneuvers. Right: Diagrammatic representation of the logic of the experiments, with a table depicting the results of a two-minute test-phase with the heat permanently switched off after eight minutes of training. (A) Operant color learning in flight simulator mode. Four identical vertical stripes can be rotated around the tethered fly using an electrical motor. Flies chose flight angles with respect to the stripes using their yaw-torque. Flight directions denoted by two opposing stripes lead to one coloration of the fly's environment (arena), flight directions towards the other two stripes to a different color-ation (i.e., blue vs. green). One of the two colors is made contiguous with heat punishment. Consecutive turning maneuvers in the same direction will rotate the arena with the stripes around the fly, into and out of the heated quadrants. Thus, no specific behavior is associated with the heat, only the coloration of the arena, leading to world-learning. (B) 'Pure' operant learning where only attempted left (or right) turning maneuvers are punished and no predictive stimuli are present. Thus, the only predictor of punishment is the behavior of the fly, leading to self-learning. (C) 'Composite' operant conditioning, where both colors and the fly's behavior are predictive of heat punishment. Left-turning yaw torque leads to one illumination of the arena (e.g., blue), while right turning yaw-torque leads to the other color (e.g., green). During training, one of these situations is associated with heat punishment. Thus, the flies have the possibility for both world- and self-learning. Interestingly, these experiments require only the biological processes known from world-learning as in (A), suggesting a hierarchical interaction between world- and self-learning. WT, Wildtype flies; cAMP, mutant flies of the strain rut^{2080} affecting a type I adenylyl cyclase deficient in synthesizing cyclic adenosine monophosphate; PKC, organism-wide downregulation of protein kinase C activity by means of an inhibitory peptide PKCi; MB, Compromised mushroom-body function be expressing tetanus neurotoxin light chain specifically in the Kenyon cells of the mushroom-bodies.
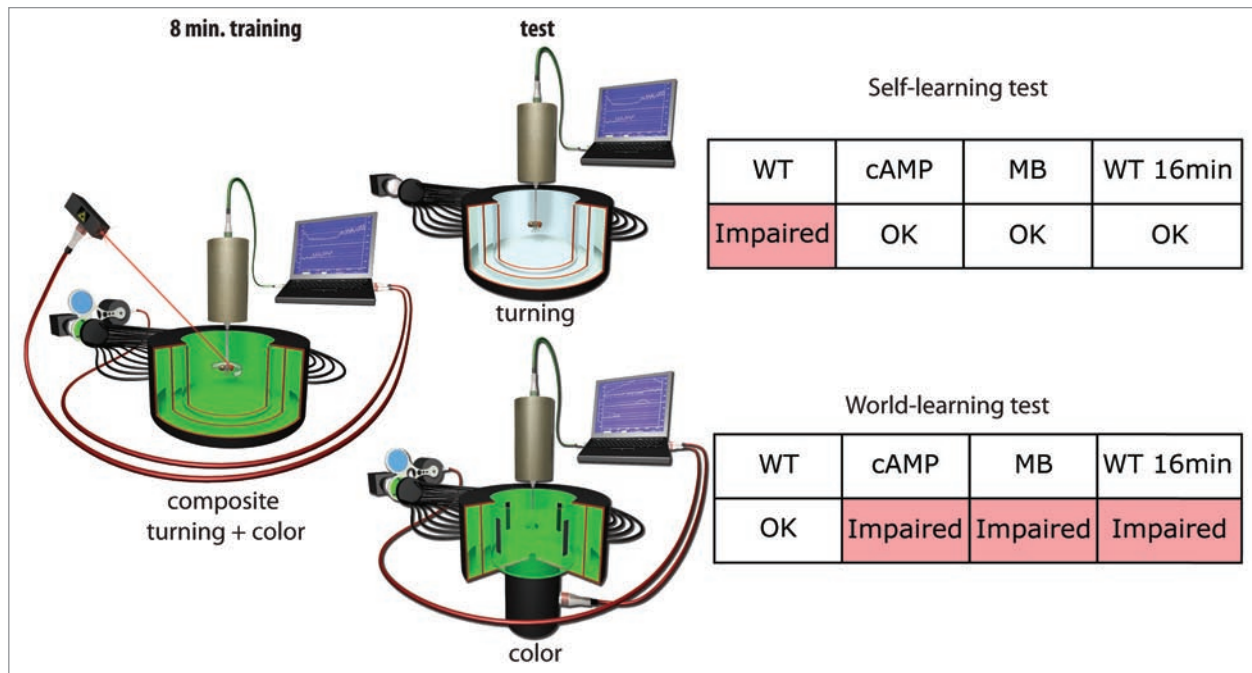
Figure 2. Isolating the two learning systems. After composite operant training (see Fig. 1C), the flies are tested either for the turning preference or for their color preference with the heat permanently switched off. Turning preference (self-learning test) is measured in a constant stimulus situation; color preference (world-learning test) is measured in the flight simulator mode described in Figure 1A. WT, Wildtype flies; cAMP, mutant flies of the strain rut^{2080} affecting a type I adenylyl cyclase; MB, Compromised mushroom-body function be expressing tetanus neurotoxin light chain specifically in the Kenyon cells of the mushroom-bodies; WT 16 min, Wildtype flies trained for 16 minutes instead of the regular eight minutes.
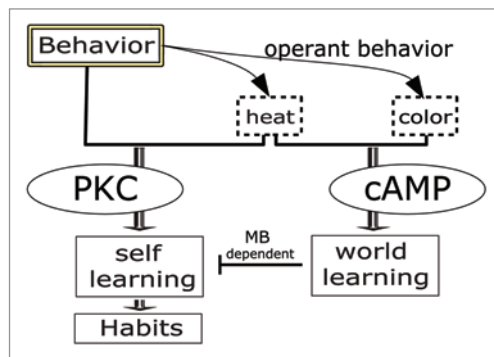
Figure 3. Conceptual model of interacting learning systems during operant conditioning. Animals use operant behavior to find out how to control sensory stimuli. If one of the stimuli carries biological value, the animal can associate this value both with other stimuli (world learning) and with its behavior (self-learning). In flies, world-learning (dependent on cAMP) inhibits self-learning (dependent on PKC) via the mushroom-bodies (MB). Extended training is required to overcome this inhibition and engage the self-learning system to form habits

6. Rescorla RA, Solomon RL. Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. Psychological Review 1967; 74:151-82.
7. Gormezano I, Tait RW. The Pavlovian analysis of instrumental conditioning. Pavlovian Journal of Biological Science 1976; 11:37-55.
8. Balleine BW, Ostlund SB. Still at the choice-Point: action selection and initiation in instrumental conditioning. Annals of the New York Academy of Sciences 2007; 1104:147-71.
9. Brembs B. Operant learning of drosophila at the torque meter. Journal of Visualized Experiments 2006–2008. 2008; doi: 10.3791/731.

10. Heisenberg M, Wolf R, Brembs B. Flexibility in a single behavioral variable of Drosophila. Learning and Memory 2001; 8:1-10.
11. Brembs B, Plendl W. Double dissociation of pkc and ac manipulations on operant and classical learning in drosophila. Current Biology 2008; 18:1168-71.
12. Davis RL. Olfactory memory formation in Drosophila: from molecular to systems neuroscience. Annu Rev Neurosci. 2005;28:275-302.
13. Isabel G, Pascual A, Preat T. Exclusive consolidated memory phases in Drosophila. Science (New York NY) 2004; 304:1024-7.
14. Lorenzetti FD, Baxter DA, Byrne JH. Molecular mechanisms underlying a cellular analog of operant reward learning. Neuron 2008; 59:815-28.

15. Nelson TJ, et al. Insulin, PKC signaling pathways and synaptic remodeling during memory storage and neuronal repair. European journal of pharmacology 2008; 585:76-87.
16. Bonini JS, et al. On the participation of hippocampal PKC in acquisition, consolidation and reconsolidation of spatial memory. Neuroscience 2007; 147:37-45.
17. Brembs B. Mushroom bodies regulate habit formation in Drosophila. Current biology: CB 2009; 19:1351-5.
18. Packard MG, McGaugh JL. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. Neurobiology of learning and memory 1996; 65:65-72.
19. Berniker M, Kording K. Estimating the sources of motor errors for adaptation and generalization. Nature neuroscience 2008; 11:1454-61.
20. Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. Nature Reviews Neuroscience 2006; 7:464-76.
21. Wolf R, Heisenberg M. Basic organization of operant behavior as revealed in Drosophila flight orientation. Journal of Comparative Physiology A, Sensory, Neural, and Behavioral Physiology 1991; 169:699-705.
23. Slamecka NJ, Graf P. Generation Effect—Delineation of a Phenomenon. Journal of Experimental Psychology: Human Learning and Memory 1978; 4:592-604.
24. Thorndike EL. Animal Intelligence. An Experimental Study of the Associative Processes in Animals. (Macmillan: New York, 1898).
25. Kornell N, Terrace HS. The Generation Effect in Monkeys. Psychological Science 2007; 18:682-5.
26. Baden-Powell R. Scouting for Boys. C. Arthur Pearson Ltd: London 1908; 288.
27. James W. The Principles of Psychology. (Holt: New York, 1890).
28. Claridge-Chang A, Roorda RD, Vrontou E, Sjulson L, Li H, Hirsh J, et al. Writing memories with light-addressable reinforcement circuitry. Cell. 2009 Oct 16;139(2):405-415