



Volume 87, No. 1, May 2011

ISSN 0376-6357
BPRODA 87 (1) 1-164 (2011)

Behavioural Processes

SQAB



Society for the Quantitative Analyses of Behavior

Special issue:

*Choice, Discounting, and the Future Value
of Quantitative Analyses*

*Guest Editors: Timothy A. Shahan, Utah State University, U.S.A.
K. Geoffrey White, University of Otago, New Zealand*



Spontaneous decisions and operant conditioning in fruit flies

Björn Brembs*

Freie Universität Berlin, Institute for Biology – Neurobiology, Königin-Luise-Strasse 28/30, 14195 Berlin, Germany

ARTICLE INFO

Article history:

Received 26 November 2010
Received in revised form 27 January 2011
Accepted 14 February 2011

Keywords:

Operant
Self-learning
World-learning
Drosophila
Insect
Learning
Memory
Multiple memory systems

ABSTRACT

Already in the 1930s Skinner, Konorski and colleagues debated the commonalities, differences and interactions among the processes underlying what was then known as “conditioned reflexes type I and II”, but which is today more well-known as classical (Pavlovian) and operant (instrumental) conditioning. Subsequent decades of research have confirmed that the interactions between the various learning systems engaged during operant conditioning are complex and difficult to disentangle. Today, modern neurobiological tools allow us to dissect the biological processes underlying operant conditioning and study their interactions. These processes include initiating spontaneous behavioral variability, world-learning and self-learning. The data suggest that behavioral variability is generated actively by the brain, rather than as a by-product of a complex, noisy input–output system. The function of this variability, in part, is to detect how the environment responds to such actions. World-learning denotes the biological process by which value is assigned to environmental stimuli. Self-learning is the biological process which assigns value to a specific action or movement. In an operant learning situation using visual stimuli for flies, world-learning inhibits self-learning via a prominent neuropil region, the mushroom-bodies. Only extended training can overcome this inhibition and lead to habit formation by engaging the self-learning mechanism. Self-learning transforms spontaneous, flexible actions into stereotyped, habitual responses.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction: operant and classical conditioning

Evolution is a competitive business. This competition has shaped the behavior of all ambulatory organisms to provide them with much more flexibility and creativity than the common stimulus–response cliché would allow them (Brembs, 2009a). In the wild, animals face a world that constantly challenges them with physically superior competitors, ever faster prey, ever more cunning predators, unpredictable weather, foreign habitats and a myriad of other, potentially life-threatening problems. In order to survive and procreate, animals have evolved not only to learn about the relationships between objects and events in this world (often studied experimentally using classical or Pavlovian conditioning), but also about how the world responds to their actions (often studied experimentally using operant or instrumental conditioning). Traditionally, both learning processes have been conceptualized as the detection and memorization of temporal contingencies, in the former case among external stimuli and in the latter between actions and external stimuli. However, most learning situations comprise both contingencies in an inextricable loop: the behaving

animal constantly receives a stream of sensory input that is both dependent and independent of its own behavior. It was the genius of Pavlov to prevent his dogs from entering this loop with the world, isolating the conditioned and the unconditioned stimulus from the control of the animal. On the face of it, Skinner’s analogous genius was to isolate the instrumental action and study the rules by which it controls its consequences. However, as the scholars at the time were well aware, the levers in Skinner’s boxes signaled food for the rats pressing them just as accurately as Pavlov’s bell signaled food for his dogs. Therefore, a recurrent concern in learning and memory research has been the question whether a common formalism can be derived for operant and classical conditioning or whether they constitute an amalgamation of fundamentally different processes (Skinner, 1935, 1937; Konorski and Miller, 1937a,b; Guthrie, 1952; Sheffield, 1965; Rescorla and Solomon, 1967; Trapold and Winokur, 1967; Trapold and Overmier, 1972; Hellige and Grant, 1974; Gormezano and Tait, 1976; Donahoe et al., 1993; Donahoe, 1997; Brembs and Heisenberg, 2000; Brembs et al., 2002; Balleine and Ostlund, 2007).

In this article I would like to review some of the new evidence for and against a hypothesis that there may be two fundamental mechanisms of plasticity, one which modifies specific synapses and is engaged by learning about the world, and one which modifies entire neurons and is engaged whenever neural circuits controlling behavior need to be adjusted. Both of these mechanisms appear to

* Corresponding author. Tel.: +49 30 838 55050; fax: +49 30 838 55455.
E-mail address: bjoern@brembs.net
URL: <http://brembs.net>.

be deeply conserved on the genetic level among all bilaterian animals. I will refer to these mechanisms as world- and self-learning, respectively, when presenting some of this evidence. It is important to emphasize that while these biological processes may to some extent be differentially recruited during certain, specific operant and classical conditioning experiments in the laboratory, they probably are engaged, to varying degrees, in many different conditioning situations. Thus, while the terms ‘classical’ and ‘operant’ are procedural definitions denoting how animals learn, self- and world-learning denote the biological processes underlying what is being learned during operant, classical or other learning situations (Colomb and Brembs, 2010).

2. Striving to emulate Pavlov: isolating the operant behavior

Because rats need a lever to press and thus may learn about the food-predicting properties of the lever, this experiment is not ideal for studying the neurobiology underlying operant learning processes. Any memory trace found in the brain cannot be unambiguously attributed to the mechanism engaged when learning about the lever or to the learning about the behavior required to press the lever. Therefore, preparations had to be developed without such ‘contamination’. One such preparation is tethered *Drosophila* at the torque meter (Heisenberg and Wolf, 1984; Wolf and Heisenberg, 1986, 1991; Wolf et al., 1992; Heisenberg, 1994; Heisenberg et al., 2001; Brembs, 2009a). For this experiment, the fly is fitted with a small hook, glued between head and thorax (Brembs, 2008). With this hook, the fly is attached to a measuring device that measures the angular momentum the fly exerts when it attempts to rotate around its vertical body axis (yaw torque; Fig. 1)(Götz, 1964; Heisenberg and Wolf, 1984). Even in the absence of any change in their sensory input, flies tethered at the torque meter show a striking variability in their yaw torque behavior (Fig. 1b)(Heisenberg, 1994). On the face of it, one may assume that this variability is mainly due to noise, as there are no cues prompting each change in turning direction. However, a mathematical analysis excluded noise as a primary cause behind the variability and instead revealed a nonlinear signature in the temporal structure of the behavior. If the fly is not changing turning directions at random and given the propensity of nonlinear system to behave random-like, it is straightforward to interpret the data as evidence for a nonlinear decision-making circuit in the fly brain determining in which direction to turn when, and with how much force (Maye et al., 2007; Brembs, 2010). Apparently, even flies are capable of making spontaneous decisions in the absence of any sensory cues eliciting or informing the decision (i.e., initiating activity (Heisenberg, 1983). Conveniently, in this setup many different environmental cues can be made contingent on many different behavioral decisions in order to design experiments exploring the neurobiology of these processes in a genetically tractable model organism (Wolf and Heisenberg, 1991; Wolf et al., 1992; Heisenberg et al., 2001). For instance, the angular speed of a drum rotating around the fly centered within it can be made proportional to the fly’s yaw torque, allowing the fly to adjust ‘flight direction’ with respect to visual patterns on the inside of the drum. One can make different wavelengths of light contingent on the sign (i.e., left or right) of the yaw torque, allowing the animal to control either the coloration (e.g. green or blue) or the temperature (i.e., infrared) of its environment. Various combinations of all these possibilities have been realized and are too numerous to mention here (Wolf and Heisenberg, 1986, 1997; Ernst, 1999; Heisenberg et al., 2001; Brembs and Heisenberg, 2001; Tang et al., 2004; Liu et al., 2006; Brembs and Hempel De Ibarra, 2006; Brembs and Wiener, 2006). Important for the argument made here is the possibility to allow the spontaneous decisions to turn in one direc-

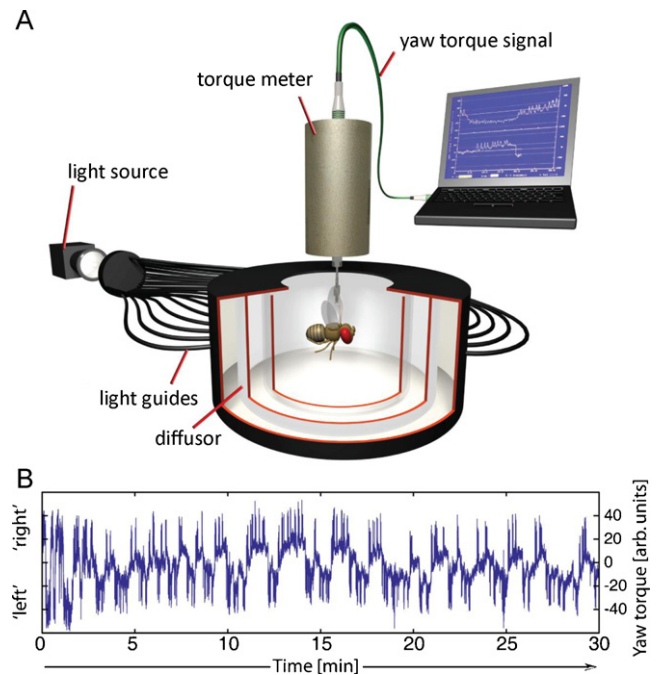


Fig. 1. Suspended at the torque meter, the fruit fly *Drosophila* initiates behavioral activity even in the absence of any change in its stimulus situation. **A** – The tethered fly is surrounded by a cylindrical drum which is illuminated homogeneously from behind (arena). The torque meter is connected to a computer recording the yaw torque traces **B** – Example yaw torque trace showing 30 min of uninterrupted flight in a completely featureless arena. Positive values correspond roughly to turning maneuvers which would rotate the fly to its right in free flight, whereas negative values would turn the fly to its left. The trace exhibits several components contributing to the variability, a slow baseline component and fast, superimposed torque spikes, corresponding to body-saccades in free flight. During the experiment, the fly initiates numerous torque spikes and many changes in turning direction. (Modified from Maye et al., 2007).

tion, say, right turning (positive torque values in Fig. 1b) to switch the environment from one state to another, say from green and hot to blue and cold or vice versa. This simple concept can be simplified even further: positive torque values can be made to lead to hot temperature, without any change in coloration present – the experiment is performed in constant white light: the only thing concomitant with the switch in temperature is the transition of the yaw torque value from one domain to the other, nothing in the environment of the fly changes other than temperature (Fig. 2). Because there are no external cues such as levers indicating which behavior will be rewarded/punished, this is one example of technically isolating the operant behavior to an extent previously unattained.

Parallel developments to isolate the operant behavior have been made in the sea slug *Aplysia* (Nargeot et al., 1997, 1999a,b,d, 2007, 2009; Brembs et al., 2002; Nargeot, 2002; Lorenzetti et al., 2006; Lorenzetti et al., 2008; Nargeot and Simmers, 2010). There, freely moving animals generate feeding movements in the absence of eliciting stimuli. Even the isolated buccal ganglia, which control these feeding movements in the intact animal, generate spontaneous motor patterns in the dish (‘fictive feeding’). Implanted electrodes can then be used to provide the animal or the isolated ganglia with a virtual food reward as reinforcement for one class of feeding movements but not others. Analogous to the Pavlovian strategy of isolating the relevant stimuli and then tracing their pathway into the nervous system until the synapse of convergence between conditioned and unconditioned stimulus had been identified, we identified the neuron where spontaneous behavior and reinforcement converge: a neuron called B51 (Plummer and Kirk, 1990; Nargeot et al., 1999b,c; Brembs et al., 2002). Impor-

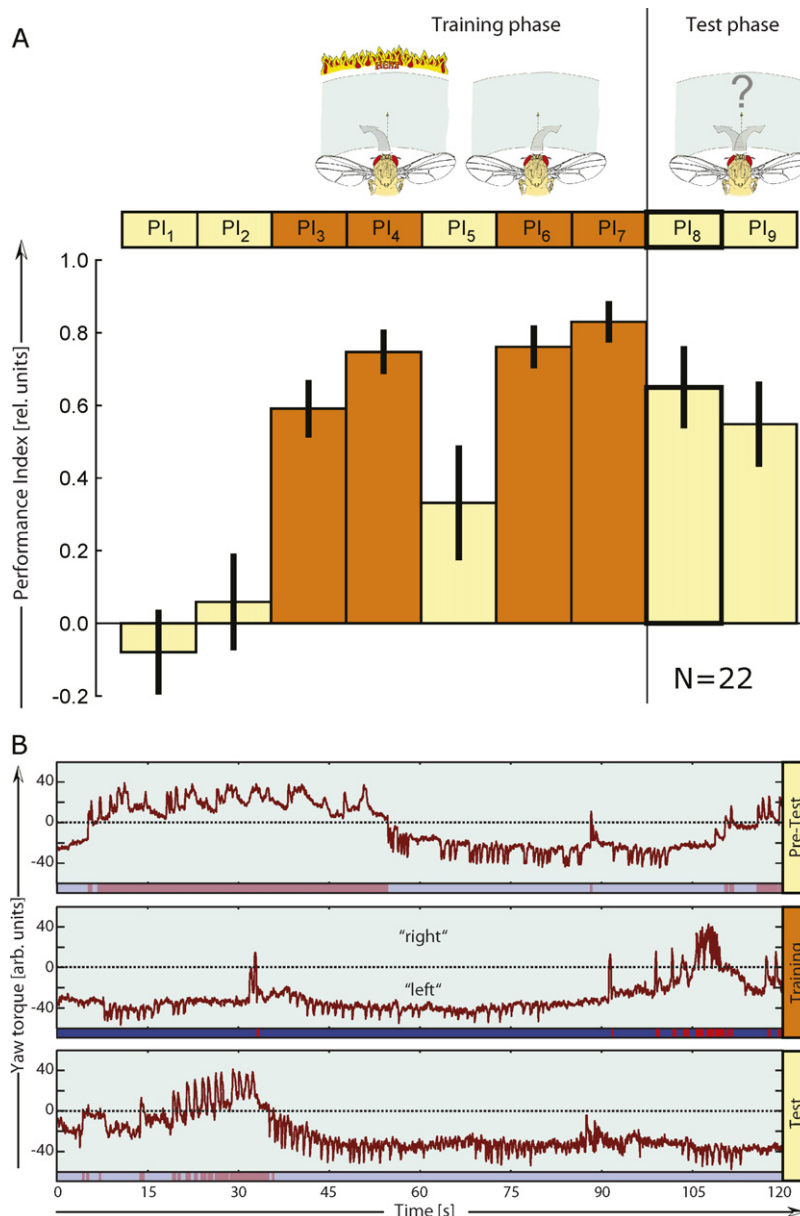


Fig. 2. Isolating the operant behavior. **A** – The layout of the experiment. The pictograms illustrate the situation for the fly (drawings generously provided by Reinhard Wolf). In this case, during the training phase, left turning attempts are associated with heat, while right turning attempts are not. In the subsequent test phase, the spontaneous decisions of the fly are recorded. The nine adjacent boxes illustrate the temporal flow of the experiment. Each box denotes one of nine consecutive two-minute periods for which a performance index (PI) is calculated. The PI = 1 if the fly spent the entire period in the unpunished situation, PI = -1 if the fly spent the entire period in the punished situation and PI = 0 if the fly distributed its choices evenly throughout the period. Recording frequency is 20 Hz, allowing for a fine-grained computation of PIs. The diagram shows an example set of $N = 22$ experiments with wildtype flies of the strain 'Canton S'. Yellow boxes/bars indicate heat off (testing periods); orange boxes/bars indicate heat under the control of the fly (training periods). **B** – Example traces of an individual fly trained to avoid right turning attempts. The first row shows one period of pre-training, with the yaw torque trace in dark red. Underneath the torque trace is a bar in unsaturated red/blue coloration indicating which turning attempts will later be punished (red) or unpunished (blue). The difference of the red and blue episodes over the sum of both is then used to calculate the PI of this period. The second row shows the last training period (PI₇) with the red/blue bar underneath indicating when the fly has actually been heated. Note that the fly keeps probing the punished turning attempts. The fly continues these forays into the punished yaw torque domain also when the heat is permanently switched off (PI₈, third row), but spends most of the time in the unpunished domain. The unsaturated red/blue bar underneath the trace indicates the episodes in which the fly would have been punished had the heat been switched on. Data from the first test period after the last training period is used to assess learning performance in the Fig. 3.

tantly, we discovered that the mechanism of plasticity in B51, unlike the synaptic plasticity found in Pavlovian learning situations, increased the excitability of the entire neuron (Brembs et al., 2002; Mozzachiodi and Byrne, 2010). This finding makes intuitive sense, because neuron B51 is only recruited into the active feeding circuitry during the behaviors we rewarded and not during other feeding movements (Nargeot et al., 1999a). Later, other neurons in this same circuit have been identified which also change their biophysical membrane properties after operant conditioning,

serving to increase the number of behaviors generated (Nargeot et al., 2009). It is tempting to generalize from this one occurrence that we may have stumbled across a dedicated mechanism which modifies behavioral circuits by altering the biophysical membrane properties of entire neurons in these circuits, biasing the whole network towards one or the other behavior by altering the probabilities with which certain network states can be attained. If this were the case, one should be able to find homologous mechanisms in other animals.

3. The neurogenetics of self-learning

Its genetic toolbox is the well-known strength of *Drosophila* as a neurobiological model system. However, without the isolation of the operant behavior from contingent external stimuli, no molecular tool would have been able to tease apart the operant/classical conundrum which has intrigued scholars for over 70 years. Making a non-directional infrared heat-beam contingent on spontaneous left or right turning maneuvers (Wolf and Heisenberg, 1991) provides instantaneous punishment without any other stimuli being contingent on the behavior (analogous to the virtual food reward being contingent on feeding motor programs in *Aplysia*). Within seconds, the fly learns (by trial and error) that its turning attempts control the unpleasant heat. After 8 min of training, it biases its spontaneous decisions towards the previously unpunished turning maneuvers, even if the heat is now permanently switched off (Fig. 2). Using mutant, wildtype and transgenic animals, we discovered that the canonical, cAMP-dependent synaptic plasticity pathway was not involved in this type of learning, but manipulating protein kinase C (PKC) signaling abolished learning in this paradigm completely (Brembs and Plendl, 2008) (Fig. 3a). In the most Skinnerian sense, the animal modifies its nonlinear decision-making circuitry such that the probability of initiating the unpunished behavior increases.

These results are mirror-symmetric compared to an experiment which is almost identical, albeit with one small difference. The small change in this experiment is that whenever the direction of turning maneuvers changes, the entire visual field of the fly instantaneously turns from one color (say, green) to another (e.g., blue). Because now the colors change both with the yaw torque and the heat, the fly has the option to learn that one of the colors signals heat, analogous to how rats in Skinner-boxes may learn that the depressed lever signals food (and the undepressed lever no food). This situation now requires the cAMP cascade of synaptic plasticity and is independent of any PKC signaling (Fig. 3b). To solve this situation, it is sufficient for the flies to learn that one of the colors is associated with the heat and then use whatever means necessary to avoid this color. In the most Pavlovian sense, the flies learn the color-heat contingency independently of the behavior with which it was acquired, they only learn about the world around them, without leaving an observable trace that the behavioral decision-making circuitry itself has been altered, even though, of course, the entire situation is still just as operantas without the stimuli (Brembs and Heisenberg, 2000; Brembs, 2009b).

These two experiments exemplify drastically why the operant/classical debate has been going on for so many decades without any solution. Here we have two obviously operant experiments, one in which the operant behavior has been experimentally isolated and one almost identical, except that in this second experiment, the behavior is not isolated any more. Instead, it is accompanied by an explicitly predictive stimulus. Both experiments are clearly operant conditioning experiments, but the biological processes cleanly separate the experiments into two different operant categories. Perhaps one factor hampering research in the previous decades was the lack of a proper terminology, which separates between procedures and mechanisms, allowing for the necessary experiments to be designed? Our discovery demonstrates that the procedural distinction between operant and classical conditioning is not helpful for understanding the neurobiological mechanisms engaged during the learning tasks. To distinguish the biological process which modifies the behavioral circuits from the one which modifies sensory pathways, we have suggested to call the first process 'self-learning' (the animal learns about its own behavior) and the second process 'world-learning' (the animal learns about the properties of the world around it) (Colomb and Brembs, 2010). Evidence from *Aplysia* suggests that this distinction is highly conserved among bilate-

rian animals: also in the neuron B51, the type I adenylyl cyclase involved in synaptic plasticity is not required to show the change in excitability after self-learning, whereas PKC signaling is necessary (Lorenzetti et al., 2008).

Another piece of evidence, as yet not published in a peer-reviewed journal, ties these insights to analogous vertebrate learning mechanisms. Vocal learning, be it human language or bird-song, follows the same operant trial-and-error principle as both our paradigms in *Drosophila* and *Aplysia*: spontaneous, variable behavior is first generated and then the sensory feedback (in this case auditory) received from this behavior is evaluated. The consequence of this evaluation is a lasting modification of the circuit generating the behavior leading to a bias towards the rewarded and away from the unrewarded behaviors. The only difference between vocal learning and invertebrate self-learning is the primary reward/punishment in the invertebrate case, compared to the more implicit reward of speaking a word or matching the tutor song. One prominent gene which has been discovered studying the biological basis of vocal learning is the gene FOXP2 (Lai et al., 2001; Fisher et al., 2009). Every member of the now famous KE family who suffers from severe verbal dyspraxia carries a mutated FOXP2 allele (Lai et al., 2001). If FOXP2 is knocked down in the basal ganglia of zebra finches, they fail to learn their song properly (Haesler et al., 2007). *Drosophila* and other invertebrate genomes also contain a *FoxP* gene (Santos et al., 2010). A mutation or RNAi-mediated knockdown of *FoxP* in *Drosophila* yields a phenocopy of our PKC manipulations, i.e., impaired self-learning and unaffected world-learning (Brembs et al., 2010; Colomb et al., 2010).

It is tempting to interpret these results as corroborating evidence for the hypothesis that there exists a highly conserved mechanism of plasticity which evolved specifically to modulate the excitability of neurons involved in behavioral choice. This mechanism appears to be based on PKC and *FoxP* signaling mediating changes to the biophysical membrane properties affecting the excitability of the entire neuron. These excitability changes, in turn, serve to bias the decision-making network towards rewarded and away from punished behaviors by making it less likely for the network to reach a state which would generate a punished behavior and making it more likely to reach a state which would generate a rewarded behavior.

4. Multiple memory systems interacting

One may wonder why synaptic plasticity was not sufficient for all learning situations. Was there a specific need for a second mechanism that could not have been served equally well by synaptic plasticity? These questions can be tackled experimentally by separate tests for the two learning systems after operant conditioning with both components present (i.e., 'composite conditioning'). Specifically, we trained animals in the right/left turning paradigm with colors present and then tested them after varying amounts of training in different test situations (Fig. 3c). Corroborating our interpretation that the colors are learned independently from the behavior used to control them, flies do not show any preference for the unpunished turning direction once the colors are removed, even though the amount of training was exactly the same as without colors (Brembs, 2009b). Conversely, the flies can avoid the punished color with a novel, orthogonal behavior controlling the colors. Thus, world-learning inhibits self-learning which would take place without the colors. If the same tests are conducted after twice the regular amount of training (i.e., 16 instead of 8 min), the results are reversed: now the flies show a spontaneous preference towards the previously unpunished direction when the colors are removed and are not able to avoid the previously punished color any more, using the orthogonal behavior. Mimicking the

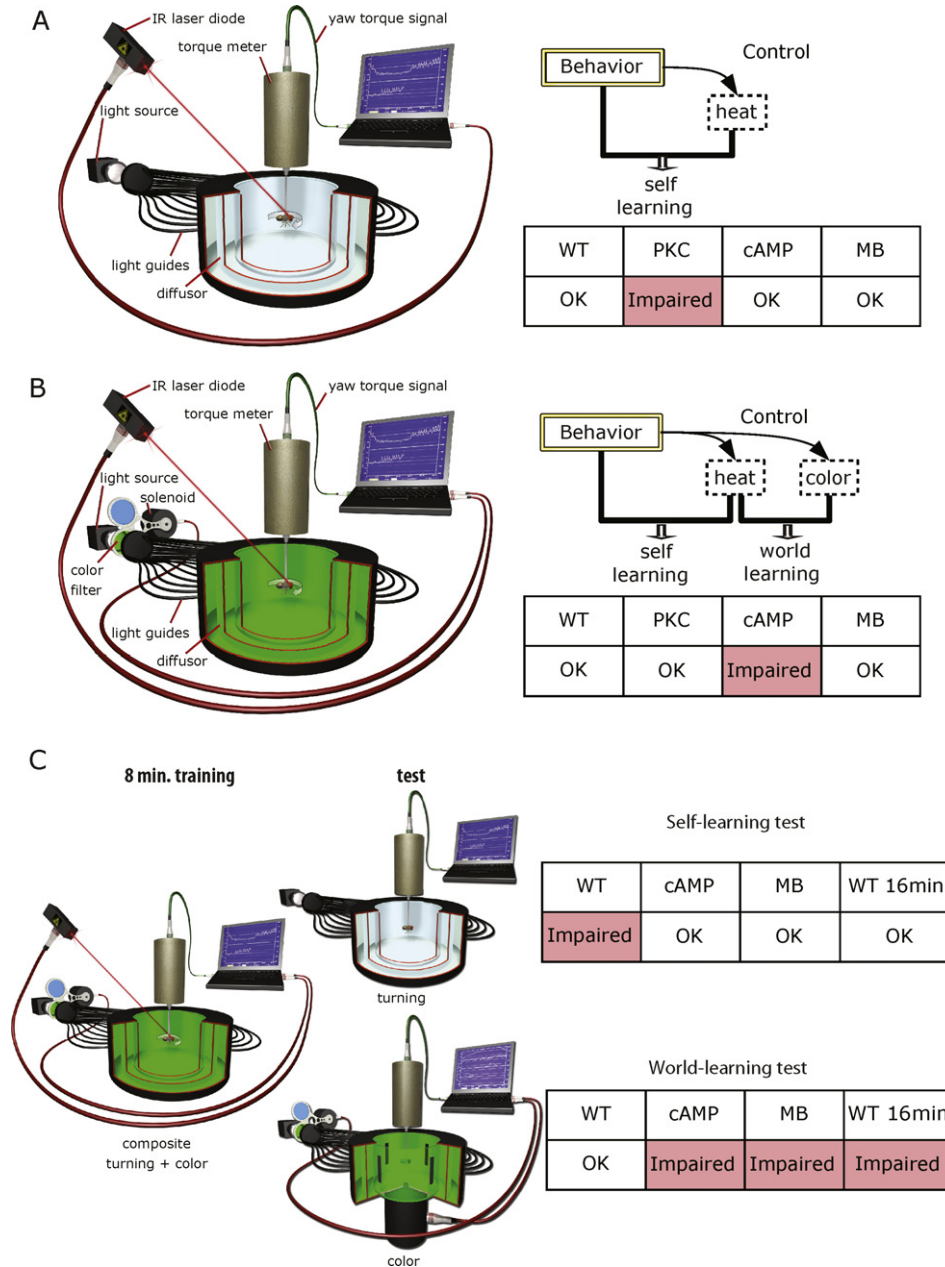


Fig. 3. Experimental procedures used to study the biological processes engaged during operant conditioning. **A** – ‘Pure’ operant learning where only attempted left (or right) turning maneuvers are punished and no predictive stimuli are present (s. Fig. 2). Thus, the only predictor of punishment is the behavior of the fly, leading to self-learning. This learning process requires neither the cAMP cascade nor the mushroom-bodies, but does require intact PKC signaling. **B** – ‘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. **C** – After composite operant training (**B**), the flies are tested either for the turning preference or for their color preference in generalization tests. Turning preference (self-learning test) is measured in a constant stimulus situation; color preference (world-learning test) is measured in the flight simulator mode. In this test, the flies control the angular position of four identical vertical stripes on the arena wall. Flies chose flight angles with respect to the stripes using their yaw-torque (right turning attempts lead to a rotation of the arena to the left and vice versa). Flight directions denoted by two opposing stripes lead to one coloration of the arena, flight directions towards the other two stripes to a different coloration (i.e., blue vs. green). Thus, the fly controls the colors using a novel, orthogonal behavior compared to the training situation. WT – Wildtype flies; cAMP – mutant flies of the strain *rut²⁰⁸⁰* 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 by expressing tetanus neurotoxin light chain specifically in the Kenyon cells of the mushroom-bodies. WT 16 min – wildtype flies trained for 16 min instead of the regular 8 min. (Modified from Colomb and Brembs, 2010).

formation of habits or skills in vertebrate animals, the extended training has overcome the inhibition of self-learning by world-learning such that world-learning could take place and modify the fly’s decision-making circuits. The flexible, goal-directed actions controlling the heat have become stereotyped, habitual responses. The scientists studying mammals commonly refer to this process as the transformation of response–outcome (R–O) associations

into stimulus–response (S–R) associations (Balleine and O’Doherty, 2010).

Thus, there is an interaction between the two learning systems which would seem difficult, if not impossible to implement if there was only one mechanism mediating all learning types. World-learning inhibits self-learning in a time-dependent manner which serves to delay modification of the behavioral circuitry until proper

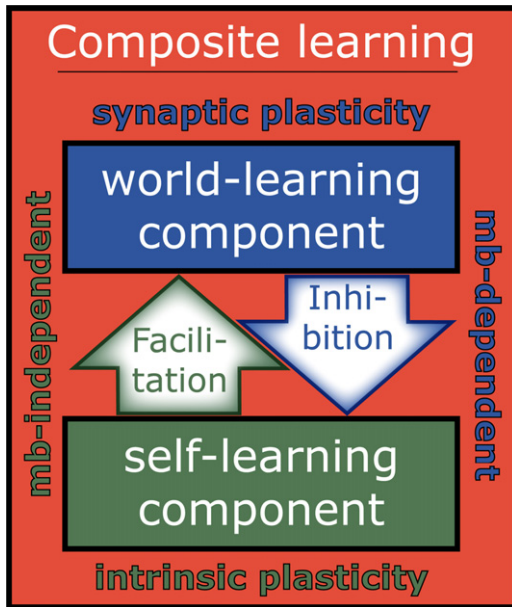


Fig. 4. Hypothetical model of composite learning consisting of two components with reciprocal, hierarchical interactions. In learning situations where the animal has the possibility to simultaneously learn about relationships between stimuli in the world and about the consequences of its own behavior, two learning systems may be engaged. One learning system learns about the stimuli (world-learning system) and the other system modifies behavioral circuits (self-learning system). The synaptic plasticity-mediated world-learning system inhibits the intrinsic plasticity-mediated self-learning via the mushroom-bodies (mb). Operant behavior controlling predictive stimuli facilitates learning about these stimuli via unknown, non-mushroom-body pathways. These interactions lead to efficient world-learning, allow for generalization and prevent premature habit-formation. (Modified from Brembs, 2008).

vetting, rehearsal or optimization has taken place. Using transgenic animals to silence certain populations of neurons in the fly brain, we discovered that this inhibition is mediated by a prominent neuropil in the insect brain, the mushroom-bodies (*corpora pedunculata*) (Brembs, 2009b).

There is yet another, reciprocal interaction between the two memory systems. This interaction facilitates world-learning when the to-be-learned stimuli are under operant control. Flies, just like virtually every animal ever tested (Thorndike, 1898; Slamecka and Graf, 1978; Kornell et al., 2007), learn a contingency in the world around them more efficiently if they can explore these contingencies themselves (Brembs and Heisenberg, 2000; Brembs and Wiener, 2006). This facility has also been termed learning by doing (Baden-Powell, 1908) or the generation effect (Slamecka and Graf, 1978).

Thus, learning in realistic situations comprising self- and world-learning components is mediated by two dedicated, evolutionary conserved learning systems which interact reciprocally. These interactions ensure efficient, yet flexible world-learning and allow for sufficient rehearsal/practice time before self-learning transforms goal-directed actions into stereotypic, habitual responses (Fig. 4).

5. Conflicting evidence

Clearly, I have painted a simplified and exaggerated hypothesis in an attempt to explain the concepts more clearly and facilitate comprehension of the biological results against the novel theoretical backdrop. Just as synaptic plasticity is rarely, if ever mediated exclusively by pre- or by postsynaptic mechanisms, exclusive recruitment of synaptic or intrinsic plasticity by any given experimental procedure will probably be equally rare. In this section, I

will provide a brief summary of the currently available evidence which is difficult to reconcile with the hypothesis fleshed out so far.

Probably the most tentative aspect of the hypothesis is the conjecture that because there is a double dissociation between *rutabaga* (i.e., type I adenylyl cyclase-mediated cAMP cascades) and PKC manipulations on world- and self-learning, respectively, in *Drosophila* and in *Aplysia*, that the same double dissociation also holds for the proposed physiological mechanisms, namely synaptic and intrinsic plasticity, respectively. The main evidence for this conjecture stems from a single model system, *Aplysia*. In *Drosophila*, we do not know which form of plasticity mediates self-learning. We only know that whatever plasticity mediates self-learning, it appears to require the same molecular cascades as the intrinsic plasticity discovered in *Aplysia*. In addition to this deduction by analogy, there is a plausibility argument: if network states are defined by the neurons participating in any given network dynamics, intrinsic plasticity is a very plausible way in which such recruitment (and hence network states) can be biased. While behavior is known to be controlled by the state of the network controlling it, sensory input is often depicted as a 'labeled line' along which the sensory information is processed and transmitted. Finally, the dominant paradigm in the literature by which classical learning is explained is synaptic plasticity, while there is little neurophysiological data for operant conditioning.

Challenging to our tentative postulation of a double dissociation among physiological mechanisms are mainly reports that classical conditioning situations, where largely world-learning mechanisms would be assumed to be at play, also seem to involve intrinsic plasticity (reviewed, for instance, in (Mozzachioldi and Byrne, 2010)). These examples include a type of photoreceptor in *Hermisenda* changing its biophysical membrane properties after classical conditioning (Crow and Alkon, 1980; Alkon, 1984; Farley and Alkon, 1985), neuron B51 in *Aplysia* changing its biophysical properties in the opposite direction after classical conditioning compared to operant conditioning (Lorenzetti et al., 2006), changes in the input resistance of LE sensory neurons in *in vitro* classical conditioning of the *Aplysia* siphon withdrawal reflex (Antonov et al., 2001), nonsynaptic plasticity in modulatory neurons after classical conditioning in the freshwater snail *Lymnaea* (Jones et al., 2003; Kemenes et al., 2006) and trace eyelid conditioning producing changes in the excitability of pyramidal neurons in the CA1 area of the rabbit hippocampus (Disterhoft et al., 1986; Moyer et al., 1996).

The fact that all of these examples all dwarfed in number by the publications demonstrating synaptic plasticity in the sensory pathways during classical conditioning may be explained by researchers not being aware of intrinsic plasticity and therefore not explicitly looking for it. However, it may also be that these examples are exceptions to a more common rule. Some of these examples also fail to report if the observed phenomena are really crucial to the learning process, i.e., would preventing these processes from happening during conditioning really impair the performance of the animals in the test after training? Some of the effects are only observed long after the learning has taken place and may be a process helping in stabilize the memory but contributing little to the actual content of the memory trace. Along similar lines, most of these examples also fail to rule out motivational or contextual contributions of the observed modifications, rather than being specific components of a memory trace. All of the above arguments serve to make the point that there probably are many more than just two learning mechanisms engaged during the various forms of conditioning and these examples may eventually provide avenues for discovering them. Thus, while publications like the ones just cited do pose problems to our hypothesis, at this point they do not suffice to eliminate it conclusively.

A second, related class of observations also poses problems for our hypothesis. This is the rare case of observations of synaptic plasticity in self-learning situations, such as the enhanced strength of electrical synapses in the central pattern generator in the *Aplysiabuccal* ganglia after operant conditioning (Nargeot et al., 2009). Indeed, this is exactly the kind of synaptic modification one would postulate to be sufficient for all kinds of learning (see above). However, the authors themselves conclude that the changes in biophysical membrane properties may be the source for the increase in electrical coupling: “it is likely that the measured increase in membrane input resistance accounts in large part for their increased coupling coefficients, with perhaps little or no additional contribution made by a direct alteration in junctional resistance” (Nargeot et al., 2009). There are also some reports on synaptic changes after operant conditioning in lever-pressing vertebrates (Jaffard and Jeantet, 1981; Kokarotseva et al., 2009). However, as these experiments still include the lever, it is impossible to say if these changes are not triggered by learning about the lever, rather than modifying the decision-making circuitry.

Finally, objections may be raised with regard to the generality of PKC being engaged in self-learning and not in world-learning. Indeed, there are many publications showing clearly that PKC mediates synaptic plasticity in many different preparations (Olds et al., 1989; Olds and Alkon, 1991; Byrne and Kandel, 1996; Drier et al., 2002; Zhao et al., 2006; Sossin, 2007; Shema et al., 2009; Villareal et al., 2009). However, many, if not all, of these publications only show that PKC activity either takes place or is required after training, during the consolidation phase. Indeed, few of these experiments explicitly compare blocking of PKC during training as opposed to blocking PKC activity after training or during retrieval only. In fact, the *Drosophila* study where this was done showed that PKC was not required during training. The requirement for PKC was for memory maintenance not acquisition. Furthermore, there are a good dozen different PKC isoforms and it is currently unknown which isoform is responsible for the self-learning effect in *Aplysia* and *Drosophila*. Thus, even if one isoform of PKC would be discovered that is critically involved in world-learning, this may not be the same isoform that appears to be involved in self-learning.

In summary, while there are isolated publications in the literature which require *ad hoc* arguments, these reports alone do not seem sufficient to falsify our, admittedly still tentative and highly simplified hypothesis. Most likely, future research will elucidate which other biological processes are engaged during the various conditioning experiments and how they interact with self- and world-learning. At this point, the research on the biological processes of self-learning is still in its infancy and postulating a physiological mechanism to go with the genetic evidence may just trigger the necessary experiments to falsify the hypothesis.

6. Epilogue

Using modern neurobiological model systems with sophisticated behavioral testing has allowed us to make significant inroads into classic behavior-analytic problems which have been recognized and described already more than 70 years ago. Despite the evidence still being patchy and comprising only few genes in a small number of model systems, overall the data seem to fit an emerging picture that makes ecological sense. The new tools and methods have only just begun to open an avenue neither Pavlov nor Skinner could dream of. About a century after Pavlov started to study behavior using physiological and analytical approaches, we have now started to gather the first glimpses into the neurobiological mechanisms underlying the probably most fundamental brain function: how to balance internal processing with external demands in order to achieve adaptive behavioral choice. These glimpses could only

be obtained using multiple model systems, each with their own, unique advantages. The time for invertebrates in neuroscience has not ended – it is just about to take off.

References

- Alkon, D.L., 1984. Calcium-mediated reduction of ionic currents: a biophysical memory trace. *Science* 226, 1037–1045.
- Antonov, I., Antonova, I., Kandel, Eric, R., Hawkins, R.D., 2001. The contribution of activity-dependent synaptic plasticity to classical conditioning. *J. Neurosci.* 21, 6413–6422.
- Baden-Powell, R., 1908. *Scouting for Boys*. C. Arthur Pearson Ltd, London.
- Balleine, B.W., Ostlund, S.B., 2007. Still at the choice-point: action selection and initiation in instrumental conditioning. *Ann. N.Y. Acad. Sci.* 1104, 147–171.
- Balleine, B.W., O'Doherty, J.P., 2010. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 35, 48–69.
- Brembs, B., 2008. Operant learning of *Drosophila* at the torque meter. *J. Vis. Exp.* 16, <http://www.jove.com/details.asp?id=731> doi:10.3791/731.
- Brembs, B., 2009a. Mushroom bodies regulate habit formation in *Drosophila*. *Curr. Biol.* 19, 1351–1355.
- Brembs, B., 2009b. The importance of being active. *J. Neurogen.* 23, 120–126.
- Brembs, B., 2010. Towards a scientific concept of free will as a biological trait: spontaneous actions and decision-making in invertebrates. *Proc. R. Soc. B* 278 (1707), 930–939.
- Brembs, B., Heisenberg, M., 2001. Conditioning with compound stimuli in *Drosophila melanogaster* in the flight simulator. *J. Exp. Biol.* 204, 2849–2859.
- Brembs, B., Heisenberg, M., 2000. The operant and the classical in conditioned orientation of *Drosophila melanogaster* at the flight simulator. *Learn. Mem.* 7, 104–115.
- Brembs, B., Hempel De Ibarra, N., 2006. Different parameters support generalization and discrimination learning in *Drosophila* at the flight simulator. *Learn. Mem.* 13, 629–637.
- Brembs, B., Lorenzetti, F.D., Reyes, F.D., Baxter, D.A., Byrne, J.H., 2002. Operant reward learning in *Aplysia*: neuronal correlates and mechanisms. *Science* 296, 1706–1709.
- Brembs, B., Pauly, D., Schade, R., Mendoza, E., Pflüger, H.-J., Rybak, J., Scharff, C., Zars, Troy, 2010. The *Drosophila* FoxP gene is necessary for operant self-learning: Implications for the evolutionary origins of language. *In Soc. Neurosci.*, p. 704.7, San Diego, Ca, USA.
- Brembs, B., Plendl, W., 2008. Double dissociation of *pkc* and *ac* manipulations on operant and classical learning in *Drosophila*. *Curr. Biol.* 18, 1168–1171.
- Brembs, B., Wiener, J., 2006. Context and occasion setting in *drosophila* visual learning. *Learn. Mem.* 13, 618–628.
- Byrne, J.H., Kandel, E.R., 1996. Presynaptic facilitation revisited: state and time dependence. *J. Neurosci.* 16, 425–435.
- Colomb, J., Brembs, B., 2010. The biology of psychology: “simple” conditioning? *Comm. Int. Biol.* 3, 142–145.
- Colomb, J., Mendoza, E., Pauly, D., Raja, S., Brembs, B., 2010. The what and where of operant self-learning mechanisms in *Drosophila*. *In: 9th International Congress of Neuroethology*, Salamanca, Spain, p. P279.
- Crow, T.J., Alkon, D.L., 1980. Associative behavioral modification in *Hermissenda*: cellular correlates. *Science* 209, 412–414.
- Disterhoft, J.F., Coulter, D.A., Alkon, D.L., 1986. Conditioning-specific membrane changes of rabbit hippocampal neurons measured in vitro. *Proc. Nat. Acad. Sci.* 83, 2733–2737.
- Donahoe, J.W., 1997. Selection networks: simulation of plasticity through reinforcement learning. *In: Packard Dorsel, V. (Ed.), Advances in Psychology*. Elsevier Science B V, pp. 336–357.
- Donahoe, J.W., Burgos, J.E., Palmer, D.C., 1993. A selectionist approach to reinforcement. *J. Exp. Anal. Behav.* 60, 17–40.
- Drier, E.A., Tello, M.K., Cowan, M., Wu, P., Blace, N., Sacktor, Todd Charlton, Yin, J.C.P., 2002. Memory enhancement and formation by atypical PKM activity in *Drosophila melanogaster*. *Nat. Neurosci.* 5, 316–324.
- Ernst, R., Heisenberg, M., 1999. The memory template in *Drosophila* pattern vision at the flight simulator. *Vis. Res.* 39, 3920–3933.
- Farley, J., Alkon, D.L., 1985. Cellular mechanisms of learning, memory, and information storage. *Ann. Rev. Psych.* 36, 419–494.
- Fisher, Simon, E., Scharff, C., 2009. FOXP2 as a molecular window into speech and language. *Trends Genet.* 25, 166–177.
- Gormezano, I., Tait, R.W., 1976. The Pavlovian analysis of instrumental conditioning. *Integr. Psych. Behav. Sci.* 11, 37–55.
- Guthrie, E.R., 1952. *The Psychology of Learning*. Harper, New York.
- Götz, K.G., 1964. Optomotorische Untersuchung des visuellen systems einiger Augenmutanten der Fruchtfliege *Drosophila*. *Kybernetik*, pp. 77–92.
- Haesler, S., Rochefort, C., Georgi, B., Licznarski, P., Osten, P., Scharff, C., 2007. Incomplete and inaccurate vocal imitation after knockdown of *foxp2* in songbird basal ganglia nucleus area X. *PLoS Biol.* 5, 12.
- Heisenberg, M., 1994. Voluntariness (Willkürfähigkeit) and the general organization of behavior. *L. Sci. Res. Rep.* 55, 147–156.
- Heisenberg, M., Wolf, R., 1984. Vision in *Drosophila*. *In: Genetics of Microbehavior*. Springer, Berlin, Heidelberg, New York, Tokio.

- Heisenberg, M., Wolf, R., Brembs, B., 2001. Flexibility in a single behavioral variable of *Drosophila*. *Learn. Mem.* 8, 1–10.
- Heisenberg, M., 1983. Initiale Aktivität und Willkürverhalten bei Tieren. *Naturwissenschaften* 70, 70–78.
- Hellige, J.B., Grant, D.A., 1974. Eyelid conditioning performance when the mode of reinforcement is changed from classical to instrumental avoidance and vice versa. *J. Exp. Psych.* 102, 710–719.
- Jaffard, R., Jeantet, Y., 1981. Posttraining changes in excitability of the commissural path-CA1 pyramidal cell synapse in the hippocampus of mice. *Brain Res.* 220, 167–172.
- Jones, N.G., Kemenes, I., Kemenes, G., Benjamin, P.R., 2003. A persistent cellular change in a single modulatory neuron contributes to associative long-term memory. *Curr. Biol.* 13, 1064–1069.
- Kemenes, I., Straub, V.A., Nikitin, E.S., Staras, K., O'Shea, M., Kemenes, G., Benjamin, P.R., 2006. Role of delayed nonsynaptic neuronal plasticity in long-term associative memory. *Curr. Biol.* 16, 1269–1279.
- Kokarovtseva, L., Jaciw-Zurakiwsky, T., MendizabalArbocco, R., Frantseva, M.V., Perez Velazquez, J.L., 2009. Excitability and gap junction-mediated mechanisms in nucleus accumbens regulate self-stimulation reward in rats. *Neuroscience* 159, 1257–1263.
- Konorski, J., Miller, S., 1937a. On two types of conditioned reflex. *J. Gen. Psych.* 16, 264–272.
- Konorski, J., Miller, S., 1937b. Further remarks on two types of conditioned reflex. *J. Gen. Psych.* 17, 405–407.
- Kornell and Terrace, Kornell, N., Terrace, H.S. 2007, The generation effect in monkeys. *Psych. Sci.* 18, 682–685. <http://www.blackwell-synergy.com/doi/abs/10.1111/j.1467-9280.2007.01959.x>.
- Lai, C.S., Fisher, S.E., Hurst, J.A., Vargha-Khadem, F., Monaco, A.P., 2001. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413, 519–523.
- Liu, G., Seiler, H., Wen, A., Zars, T., Ito, K., Wolf, R., Heisenberg, M., Liu, L., 2006. Distinct memory traces for two visual features in the *Drosophila* brain. *Nature* 439, 551–556.
- Lorenzetti, F.D., Mozzachiodi, R., Baxter, D.A., Byrne, J.H., 2006. Classical and operant conditioning differentially modify the intrinsic properties of an identified neuron. *Nat. Neurosci.* 9, 17–29.
- Lorenzetti, F.D., Baxter, D.A., Byrne, J.H., 2008. Molecular mechanisms underlying a cellular analog of operant reward learning. *Neuron* 59, 815–828.
- Maye, A., Hsieh, C.-H., Sugihara, G., Brembs, B., 2007. Order in spontaneous behavior. *PLoS ONE* 2, e443.
- Moyer, J.R., Thompson, L.T., Disterhoft, J.F., 1996. Trace eyeblink conditioning increases CA1 excitability in a transient and learning-specific manner. *J. Neurosci.* 16, 5536–5546.
- Mozzachiodi, R., Byrne, J.H., 2010. More than synaptic plasticity: role of nonsynaptic plasticity in learning and memory. *Trends Neurosci.* 33, 17–26.
- Nargeot, R., 2002. Correlation between activity in neuron B52 and two features of fictive feeding in *Aplysia*. *Neurosci. Lett.* 328, 85–88.
- Nargeot, R., Baxter, D.A., Byrne, J.H., 1997. Contingent-dependent enhancement of rhythmic motor patterns: an in vitro analog of operant conditioning. *J. Neurosci.* 17, 8093–8105.
- Nargeot, R., Baxter, D.A., Byrne, J.H., 1999a. In vitro analog of operant conditioning in *Aplysia* II. Modifications of the functional dynamics of an identified neuron contribute to motor pattern selection. *J. Neurosci.* 19, 2261–2272.
- Nargeot, R., Baxter, D.A., Patterson, G.W., Byrne, J.H., 1999b. Dopaminergic synapses mediate neuronal changes in an analogue of operant conditioning. *J. Neurophys.* 81, 1983–1987.
- Nargeot, R., Baxter, D.A., Byrne, J.H., 1999c. In vitro analog of operant conditioning in *Aplysia* I. Contingent reinforcement modifies the functional dynamics of an identified neuron. *J. Neurosci.* 19, 2247–2260.
- Nargeot, R., Baxter, D.A., Byrne, J.H., 1999d. In vitro analog of operant conditioning in *Aplysia* II. Modifications of the functional dynamics of an identified neuron contribute to motor pattern selection. *J. Neurosci.* 19, 2261–2272.
- Nargeot, R., Le Bon-Jego, M., Simmers, J., 2009. Cellular and network mechanisms of operant learning-induced compulsive behavior in *Aplysia*. *Curr. Biol.* 19, 975–984.
- Nargeot, R., Pettrissans, C., Simmers, J., 2007. Behavioral and in vitro correlates of compulsive-like food seeking induced by operant conditioning in *Aplysia*. *J. Neurosci.* 27, 8059–8070.
- Nargeot, R., Simmers, J., 2010. Neural mechanisms of operant conditioning and learning-induced behavioral plasticity in *Aplysia*. *Cell. Mol. Life Sci.*, doi:10.1007/s00018-010-r0570-9.
- Olds, J.L., Alkon, D.L., 1991. A role for protein kinase C in associative learning. *New Biol.* 3, 27–35.
- Olds, J.L., Anderson, M.L., McPhie, D.L., Staten, L.D., Alkon, D.L., 1989. Imaging of memory-specific changes in the distribution of protein kinase C in the hippocampus. *Science* 245, 866–869.
- Plummer, M.R., Kirk, M.D., 1990. Premotor neurons B51 and B52 in the buccal ganglia of *Aplysia californica*: synaptic connections, effects on ongoing motor rhythms, and peptide modulation. *J. Neurophys.* 63, 539–558.
- Rescorla, R.A., Solomon, R.L., 1967. Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psych. Rev.* 74, 151–182.
- Santos, M.E., Athanasiadis, A., Leitão, A.B., Dupasquier, L., Sucena, E., 2010. Alternative splicing and gene duplication in the evolution of the FoxP gene sub-family. *Mol. Biol. Evol.* 28, 237–247.
- Sheffield, F.D., 1965. Relation of classical conditioning and instrumental learning. In: Prokasy, W.F. (Ed.), *Classical Conditioning*. Appleton-Century-Crofts, New York, pp. 302–322.
- Shema, R., Hazvi, S., Sacktor, Todd, C., Dudai, Y., 2009. Boundary conditions for the maintenance of memory by PKMzeta in neocortex. *Learn. Mem.* 16, 122–128.
- Skinner, B.F., 1935. Two types of conditioned reflex and a pseudo type. *J. Gen. Psych.* 12, 66–77.
- Skinner, B.F., 1937. Two types of conditioned reflex: a reply to Konorski and Miller. *J. Gen. Psych.* 16, 272–279.
- Slamecka, N.J., Graf, P., 1978. Generation Effect - Delineation of a Phenomenon. *J. Exp. Psych.: Hum. Learn. Mem.* 4, 592–604.
- Sossin, W.S., 2007. Isoform specificity of protein kinase Cs in synaptic plasticity. *Learn. Mem.* 14, 236–246.
- Tang, S.M., Wolf, R., Xu, S.P., Heisenberg, M., 2004. Visual pattern recognition in *Drosophila* is invariant for retinal position. *Science* 305, 1020–1022.
- Thorndike, E.L., 1898. *Animal Intelligence. An Experimental Study of the Associative Processes in Animals*. Macmillan, New York.
- Trapold, M.A., Overmier, J.B., 1972. The second learning process in instrumental conditioning. In *Classical conditioning II*. In: Black, A.H., Prokasy, W.F. (Eds.), *Current Research and Theory*. Appleton-Century-Crofts, New York, pp. 427–452.
- Trapold, M.A., Winokur, S., 1967. Transfer from classical conditioning and extinction to acquisition, extinction, and stimulus generalization of a positively reinforced instrumental response. *J. Exp. Psych.* 73, 517–525.
- Villareal, G., Li, Q., Cai, D., Fink, A.E., Lim, T., Bougie, J.K., Sossin, W.S., Glanzman, D.L., 2009. Role of protein kinase C in the induction and maintenance of serotonin-dependent enhancement of the glutamate response in isolated siphon motor neurons of *Aplysia californica*. *J. Neurosci.* 29, 5100–5107.
- Wolf, R., Heisenberg, M., 1986. Visual orientation in motion-blind flies is an operant behavior. *Nature* 323, 154–156.
- Wolf, R., Heisenberg, M., 1997. Visual space from visual motion: turn integration in tethered flying *Drosophila*. *Learn. Mem.* 4, 318–327.
- Wolf, R., Heisenberg, M., 1991. Basic organization of operant behavior as revealed in *Drosophila* flight orientation. *J. Comp. Physiol.* A169, 699–705.
- Wolf, R., Voss, A., Hein, S., Heisenberg, M., 1992. Can a fly ride a bicycle? Discussion on Natural and Artificial Low-Level Seeing Systems. *Phil. Trans. Roy. Soc. B: Biol. Sci.* 337, 261–269.
- Zhao, Y.L., Leal, K., Abi-Farah, C., Martin, K.C., Sossin, W.S., Klein, M., 2006. Isoform specificity of PKC translocation in living *Aplysia* sensory neurons and a role for Ca²⁺-dependent PKC APL I in the induction of intermediate-term facilitation. *J. Neurosci.* 26, 8847–8856.