Operant Behavior in Model Systems

Björn Brembs, University of Regensburg, Regensburg, Germany

Operant and Classical Conditioning

One of the traditional dichotomies in learning and memory research is that between operant (instrumental) and classical (Pavlovian) conditioning. The distinction between operant and classical conditioning is merely operational: these terms denote *how* a learning experiment was conducted. In classical conditioning, the animal's behavior has no influence over the stimuli it is presented with, while in operant conditioning the animal is in control at least of the most relevant stimuli. The stimuli and behaviors in both classes of experiments can be almost arbitrarily exchanged, as long as these rules are obeyed. Already in the 1930s, Skinner and his contemporaries realized that, in its simplest form, one could conceptualize *what* is being learned in a classical conditioning experiment in only one single process, namely the association between two stimuli. Following Pavlov, this concept can be described as "stimulus substitution" where the conditioned stimulus (CS; e.g., a tone) receives some of the response-eliciting properties of the unconditioned stimulus (US; e.g., food). Note that this does not preclude the design of less minimal experiments where the animals are allowed to learn more. Operant conditioning, on the other hand, most likely comprises at least two, if not more processes: the association between two stimuli and the association between the behavior of the animal and one or more stimuli (Skinner, 1935, 1937; Konorski and Miller, 1937a, 1937b). This is what Skinner called the "three-term-contingency" of (1) the stimuli that are present just before and during the behavior, (2) the behavior itself, and (3) the stimuli that the animal is presented with as a consequence of its behavior (Skinner, 1969). Phrased differently, it was Pavlov's strike of genius to prevent his dogs from interfering with the experiment by firmly tying them down, as it significantly reduced the number of possible learning processes to those involved in processing external stimuli.

From this brief recapitulation it becomes clear that already 80 years ago, there was very little intrinsic reason to specifically juxtapose these two types of experiments—the reason these classes of experiments are still taught and discussed together today (e.g., Domjan, 2016) are more historical than logical. In these eight decades, a wealth of literature has accumulated describing experiments in which seemingly every possible relation and detail in the configurations of stimuli and behaviors has been tested to compare and contrast operant and classical conditioning. Despite all these efforts, until the advent of modern neurobiological techniques, the more fundamental questions that Skinner and his contemporaries started tossing around in the 1930s remained largely unanswered.

Spontaneous Actions

While the relationship of the animal's behavior to antecedents was always intuitively obvious in classical conditioning—after all, the US was chosen such as to elicit a clear response—this relationship remained hotly debated in operant conditioning until rather recently. Despite emphasizing the exploratory or spontaneous nature of the ("emitted," not "elicited") behavior in operant conditioning experiments (Skinner, 1938), the behavior was nevertheless routinely referred to as a "response," implying some eliciting factor (Dickinson, 1985). This ambiguity about whether or not one needed to describe the behavior of animals in classical and operant conditioning experiments in different terms has provoked discussions until this day (Domjan, 2016). Arguably, the ambiguity may be traced to the behaviorist effort to separate behaviorism from the prevailing animist psychology of the early 20th century on one side, while at the same time emphasizing the difference to the more conventional schools of classical conditioning,

e.g., reflexology, on the other. As demonstrating spontaneous behavior would require to demonstrate the absence of relevant antecedent stimuli, making a firm case for operant behavior deserving a separate term proved elusive (Domjan, 2016).

One of the arguments in this debate can be summarized as an argument from incredulity: How could a behaving system evolve that produces behavior unrelated to its environment? How could such organisms successfully coordinate their actions with the world in which they live (e.g., Domjan, 2016)? Today, we know from various observations both in the field and in the laboratory that organisms without the capability to generate spontaneous actions would likely not have survived long. Individual behavioral variability has been found to be ecologically advantageous in game theoretical studies (McNamara et al., 2004; Glimcher, 2005, 2003; Glimcher and Rustichini, 2004; Brembs, 1996), in pursuit-evasion contests such as predator/prey interactions ("Protean Strategy") (Grobstein, 1994; Driver and Humphries, 1988; Shultz and Dunbar, 2006; Miller, 1997), in exploration/foraging (Belanger and Willis, 1996; Hills et al., 2013; Humphries and Sims, 2014; Shlesinger, 2009; Reynolds et al., 2016), in mobbing attack patterns by birds (Humphries and Driver, 1970), and in the variation of male songbirds' songs (Neuringer, 2004). Clearly, invariable behavior will be exploited (Miller, 1997; Jabloński and Strausfeld, 2000, 2001; Catania, 2009, 2010, 2008; Mitra et al., 2009; Corcoran et al., 2009) and leaves the animal helpless in unpredictable situations (Brembs, 2009b, 2010; Heisenberg, 1994).

Two particular cases out of the multitude of examples for the evolutionary benefits of spontaneous behavior deserve special mention, both because the neurobiological mechanisms underlying the behavior are particularly well understood and because these behaviors have been used in operant conditioning experiments.

Aplysia *Feeding*

The first example is that of the feeding behavior of the marine snail *Aplysia*. These animals feed primarily on different species of seaweed, the texture of which can range from mushy to too tough to eat. The animals use their radula, a tongue-like organ, to grasp food and push it into the esophagus. The radula typically consists of two halves, which can be either protracted or retracted while they are either open or closed. During initial feeding attempts, the animal protracts its radula in the open state, then closes the two radula halves and then retracts the radula, hopefully with a large chunk of edible seaweed enclosed in it (ingestion). However, if the food item is too tough to push it further into the esophagus, the animal can also protract the radula while it is still closed around the food, pushing it out of the mouth. The animal then retracts the radula in its open state (rejection).

There are two aspects to the spontaneous components of this behavior.

The first one concerns the spontaneous variability in the movement dynamics of the first attempt to grasp the food and transport it toward the gut. *Aplysia* initially generates highly variable biting and swallowing movements, in an exploratory phase of feeding. During this time, the animal is exploring the state space of the motor system in the buccal ganglia controlling the movements of the radula. The animal is trying out which movements are most effective in terms of transporting food into the esophagus (Lum et al., 2005; Horn et al., 2004). Reafferent (operant) feedback from dopaminergic fibers in the esophageal nerve instructs the buccal ganglia which movements were more or less effective (Nargeot et al., 1999c). This initial exploratory variability happens with all novel food items; it can thus be said to be independent of the nature of the food stimulus.

The second aspect concerns biting behavior in the absence of any food. Hungry *Aplysia* will spontaneously emit bites in the complete absence of any food stimulus. Clearly, it is always possible to argue that some undefined stimulus, commonly insufficient to elicit biting, becomes salient enough to trigger biting when the animal is hungry enough. However, in addition to any perceived implausibility of this argument, there is experimental evidence of neurons, which have evolved precisely to generate spontaneous behaviors. Isolated buccal ganglia of hungry *Aplysia* will continue to generate the neural programs controlling the radula in the intact animal, when placed in a Petri dish with suitable medium (Nargeot et al., 1999b, 1999a). These buccal motor programs (BMPs) occur spontaneously in the most explicit of terms: all sensory organs have been cut away; there cannot be any triggering stimuli in the Petri dish. Moreover, the timing of BMPs is not rhythmic, the type of BMPs varies between ingestion- and rejection-like patterns and the intra-BMP dynamics vary from BMP to BMP. In other words, the in vitro experiments validate the in vivo work on spontaneous, exploratory biting behavior in *Aplysia*.

Physiological work on the neurons involved in generating BMPs identified a class of neurons that appears to have evolved to generate crucial aspects of this spontaneity (reviewed in Nargeot and Simmers, 2012). Among the different classes of neurons controlling radula movements are two in particular, which are relevant here. One of them, let us call them "*What*" neurons, fires during the BMP and determine what type (ingestion- or rejection-like) of BMP will be produced by controlling the timing of activity in the closure motor neuron. A second class of neurons, let us call them "*When*" neurons, fires right before a BMP can be observed in the motor nerves. Experimentally stimulating any of these *When* neurons will lead to the recording of a consecutive BMP in the motor neurons immediately following the stimulation. Thus, while the *What* neurons determine the type of behavior, the *When* neurons determine when a given behavior is going to be emitted. Part of the variability in the intra-BMP temporal dynamics comes from each BMP being initiated by a different neuron in the *When* class. One reason for the variability in which of the *When* neurons will start a BMP lies in the weak coupling between these neurons preventing a stereotyped sequence of activity from emerging between the *When* neurons. A second crucial contribution to this spontaneous variability besides the weak coupling between them is the capability of individual *When* neurons to generate spontaneous bursts of activity even in the complete absence of input from other neurons. In contrast to the canonical neurons commonly studied, tonically stimulating an experimentally isolated *When* neuron to near its firing threshold leads to arrhythmic firing of the neuron (Nargeot and Simmers, 2012). The underlying molecular mechanisms are just being discovered and involve mathematically unstable, nonlinear calcium dynamics (Bedecarats et al., 2015). In the context of what roles these neurons play in the control of spontaneous behavior, it is straightforward to argue that these calcium dynamics have evolved to support ecologically relevant behavioral spontaneity.

In an analog of the operant feedback described earlier, one can use any of these spontaneously generated BMPs and pair them with contingent stimulation of the esophageal nerve (Nargeot et al., 1999b, 1999a). This experiment thus establishes an in vitro operant conditioning paradigm out of a behavioral observation in freely behaving animals. This fortunate combination of intact and in vitro experiments, together with the exquisite physiological accessibility of *Aplysia*, allows for an unmatched rigor in the study of the learning processes actually taking place in the nervous system during operant conditioning. Perhaps not surprisingly, one such process directly affects the spontaneity of the behavior. Further discussion on *Aplysia* feeding behavior is provided later in the chapter.

Taken together, evidence both from intact animals and deafferentated nervous systems demonstrates which aspects and components of *Aplysia* feeding behaviors are spontaneous in nature and what ultimate and proximate functions this spontaneity serves.

Drosophila *Flight*

Drosophila fruit flies with one injured wing are perfectly capable of flying straight, provided the injury is not too severe to prevent flight completely. This must seem like a remarkable feat to anyone steeped in the literature on how fixed optomotor reflexes control straight flight in *Drosophila*. Unless the animal is born with the knowledge of exactly which kind and amount of wing damage leads to which effect on torque kinematics, there must be some reafferent feedback that instructs the flies on the effectiveness of their turning maneuvers. Several observations support the conclusion that flying straight is primarily an operant behavior (reviewed in, e.g., Brembs, 2009b). For instance, one can experimentally eliminate all optomotor responses (without rendering them blind) and the flies are still able to fly straight. Crucially, these manipulated flies can do this even if the feedback between their behavior and the environment is experimentally reversed, i.e., left turning attempts lead to visual feedback that suggests a right turn and vice versa (similar to inversion goggles in humans). These experiments are done with the flies tethered to a torque meter, which measures the yaw torque of the flies without them actually rotating in space. Surrounded by a visual panorama that can be rotated around the flies instead, the experimenter has exquisite control over the flies' stimulus situation. Tethered optomotor-disabled flies manage to keep the visual panorama from rotating (i.e., fly straight) no matter how the rotation is coupled to their torque behavior. This is in striking difference to unmanipulated flies, which need a very long time until they manage to keep the panorama steady with inverted feedback (but do so within milliseconds in regular coupling). Apparently, the inborn optomotor responses impede the wild-type flies in this experiment, while the symmetrical behavior of the optomotor-impaired flies is the evidence that they must use operant behavior to minimize the amount of arena rotation in the absence of any sensory system telling them the direction of rotation (Wolf and Heisenberg, 1986). As with the example of *Aplysia* explained in previous section, with these flies being partially deafferentated, any of their turning attempts must be spontaneous as there are no optomotor stimuli being perceived that could elicit a turning response.

Another experiment brings us back to the ability of the flies to compensate for one-sided wing injury. The kind of turning maneuvers, the optomotor-impaired flies are using to control their visual feedback, can also be observed in tethered wild-type animals in a completely uniform environment. Similarly to the manipulated flies, there are no visual stimuli known to elicit turning attempts. Great care is taken to ensure that all the stimuli that are present, of any modality, are as constant as experimentally possible. In fact, from the experiments with external stimuli present, it is known that the remaining, constant stimuli do not exert any detectable effect on top of the explicit stimuli. Again, without complete experimental deafferentation, it is impossible to be sure that these sensory-deprived animals are not responding to otherwise unphysiological stimuli that bear no relation to flight under normal circumstances. One would assume that these stimuli would be occurring randomly, without any specific structure or pattern. In that case, the temporal dynamics of the attempted turning behavior of such deprived flies should be reminiscent of random noise. However, the temporal dynamics of these flies is much more reminiscent of the mathematically unstable, nonlinear process discovered in *Aplysia* (Maye et al., 2007). In a recent transgenic screen, we are beginning to identify candidate circuits comprising the neurons that may be responsible for this nonlinear signature. Interestingly, these structures appear to be located in the same brain regions as those associated with the temporal structure of walking behavior (Martin et al., 2001). Recent optophysiological work on zebra fish is also beginning to identify the brain regions in the vertebrate brain where the temporal structure spontaneous turning movements are controlled (Dunn et al., 2016).

Instead of visual feedback, one can use heat as a feedback in the tethered flight experiment and let the fly use its turning attempts to control the punishing heat beam (in a completely homogeneous visual environment). When, e.g., left turning attempts are punished, the fly shifts its baseline torque toward the unpunished (e.g., right) direction, without eliminating superimposed bidirectional fluctuations in torque, even after the heat is permanently switched off (Wolf and Heisenberg, 1991). This capability is precisely what would be required to compensate the reduced torque of a damaged wing: a shift in the baseline torque output, operantly matched to the (visual) feedback, without eliminating the possibility for left and right turns from this new "straight-flight" baseline.

Taken together, converging evidence from *Aplysia* and *Drosophila* suggests that spontaneous behavior is controlled by an evolutionary conserved mechanism that relies on unstable nonlinearities to explore the state space of the motor system in search for favorable behavioral outputs.

Behavior Is Likely Never a "Response"

Given the spontaneous components in even highly stereotyped behaviors such as feeding or optomotor control, one is tempted to examine what one may consider the simplest stimulus–response systems for evidence of spontaneity.

Phototaxis in Insects

One could consider the movement of insects toward a light source (phototaxis) as one such system. The fact that insects get trapped at windows or the proverbial moth that dies in a candle flame is an iconic example of a rigid stimulus–response organization of this behavior. But does insect phototaxis stand up to scientific scrutiny? Already a 100 years ago, observations by McEwen (1918) suggested that this behavior may be more complex than one would at first expect. McEwen found that only startled flies would walk toward the light in his small tube. Sitting flies did not seem to find light very attractive, suggesting that more than just light hitting the retina must be responsible for triggering phototaxis. He also observed that flies with clipped wings do not approach the light anymore, even when compared to walking, intact flies. This observation was later confirmed in a different experiment, where the wings were not only clipped but also genetically rendered useless for flight (Benzer, 1967). Further experimentation revealed that at least for insects, the term "phototaxis" may be inappropriate (Gorostiza et al., 2016). Rather than just affecting the approach of a light source, the flies' ability to fly affected their light/dark preference across several different behavioral tests, none of which tested phototaxis, but forced the flies to choose between more or less bright stimuli. If flying ability was compromised only temporarily, the flies' photopreference reversed concomitantly. Neuronal activity in circuits expressing dopamine and octopamine, respectively, doubly dissociated in this case of behavioral flexibility (Gorostiza et al., 2016): activity in octopaminergic neurons was necessary and sufficient to shift the flies', preference toward darkness, while activity in dopaminergic neurons was necessary and sufficient to shift the preference toward brightness. The involvement of these biogenic amines suggests that valuation of stimuli may play a role in the flies' shifts in photopreference. Apparently, flies monitor their ability to fly, and the outcome of this evaluation exerts a fundamental effect on action selection—including, but not exclusively in phototaxis experiments. This work suggests that even innate preferences that appear simple and hard wired, such as those expressed in classic phototaxis experiments, comprise a value-driven decision-making stage, negotiating external stimuli with the animal's internal state and likely other factors as well, before an action is selected. This endows the animal with the possibility to decide, for example, when it is better to move toward the light or hide in the shadows. Moreover, the fact that flies adapt their photopreference in accordance with their flying ability shows that flies have the cognitive tools required to evaluate the capability to perform an action and to let that evaluation impact other actions—an observation reminiscent of metacognition.

Thus, what appears to be a simple response to light actually contains at least one decision-making stage, which negotiates and weighs several different factors before selecting an action. One may even argue that what was commonly described as a taxis, in this case at least, only appears as a simple taxis at a superficial glance. Once one peers into the neurobiology of the behavior, the usage of terms such as "response" or "taxis" appears inadequate.

Knee Jerk Reflexes in Mammals

A similarly iconic input–output system is the "knee jerk" class of spinal reflexes. One can hardly imagine a simpler system: the 1a afferents send the excitation from the stretched muscle spindles to the γ motor neurons, which activate the muscle that flexes the leg. On the surface, this seems to be an even simpler and more obviously feedforward case than phototaxis in insects: two neurons, one synapse, input from the sensory neuron leads to output from the motor neuron and behavior. However, even there, upon closer examination, the seemingly dominant stimulus-response organization starts to collapse.

Implanting cuff electrodes on the mammalian posterior tibial nerve as well as electrodes recording the electromyograms from the soleus muscle allows for long-term recordings of the electrical analog of the stretch reflex, the H-reflex (or Hoffmann's reflex). The simple textbook case of the knee jerk reflex implies all-or-nothing responses to identical stretch stimuli. However, triggering this reflex over hours, days, or weeks reveals multiple timescales of variability in the amplitude of the H-reflex. Moreover, making a food or water reward contingent on larger (or smaller, respectively) reflex amplitudes than baseline averages leads to an increase (or decrease, respectively) of the reflex amplitude over the course of a few days (Wolpaw, 2010; Chen and Wolpaw, 1996; Wolpaw and Chen, 2006; Carp et al., 2006; Thompson and Wolpaw, 2014a, 2014b). One could say that, conceptually, this procedure is the opposite of an omission schedule: to eliminate potential operant components in classical conditioning experiments, an omission schedule leaves out the US (often food or water), whenever the conditioned reflex was produced. In operant conditioning of the H-reflex, only those reflexes are rewarded that reach the target amplitude. In this case, always the same electrical stimulus is eliciting the reflex, so the behavioral variability must be either due to internal processes in the animal or due to unrelated stimuli in the environment. It is difficult to imagine that, for instance, a certain corner of the experimental chamber consistently increases reflex amplitudes throughout the nervous system, such that the animal would always seek this corner to increase its amplitudes during conditioning. On the contrary, such operant conditioning in humans is used to improve rehabilitation after spinal cord injury (Thompson and Wolpaw, 2015, 2014a), such that lasting changes in reflex amplitude must be brought about by processes that are independent from the current environment of the patients. In fact, the function of the spontaneous variability in reflex amplitude is quite well understood: the spinal reflexes constantly adapt to the environment in which vertebrates walk. They accomplish this feat by constantly changing their amplitude and evaluating the sensory consequences of these actions. In other words, even spinal reflexes are using spontaneous actions to constantly explore the state space of the motor system to find the most suitable behavioral output. In this process, the spinal reflexes rely on networks that span the corticospinal tract all the way into the cerebellum and motor cortex. Without these reflexes using these networks to constantly probing the environment's responses to their spontaneous actions, we would not be able to walk up the stairs at the end of the hallway or climb down a mountain after we reached the summit.

Thus, the textbook knee jerk reflex only exists in its textbook form if the slice of the spinal cord that contains its synapses is removed from the rest of the nervous system. However, this isolated state prevents understanding of the function of spinal reflexes for locomotion. Describing these spinal reflexes as rigidly responding to stretch stimuli seems at best inadequate and at worst misleading, considering the large-scale networks within which this sensorimotor synapse is embedded and the complex, operant processes these circuits in fact mediate.

Phototaxis in a Polychaete Larva

A less iconic but perhaps yet simpler and likely one of the most archaic of these stimulus–response systems is also considered a model for the last common ancestor of vertebrates and invertebrates, the "Urbilaterian": the planktonic larvae of a marine polychaete worm, *Platynereis dumerilii*. These egg-shaped creatures use a band of ciliated cells around their body to locomote, in the first half of larval development preferentially toward the surface. One factor in this upward movement is positive phototaxis. When a light is switched on at one end of a chamber filled with *P. dumerilii* larvae, they all start swimming toward it (Jékely et al., 2008). That these animals are capable of phototaxis is remarkable as their nervous system does not feature any interneurons. Their light-sensitive neurons make direct synaptic connections with the ciliated cells that propel the animal. Importantly, the ciliated cells are constantly active, propelling the animals sometimes in this direction, sometimes in that. When light hits one of their two "eyes," the synaptic connection between the light-sensitive neuron and the ciliated cells inhibits the ciliated cells on the side where the light was perceived (Jékely et al., 2008). Like a rower who stops rowing on one side, the animals then rotate toward the side where the light came from. If they keep rotating, light will hit their other eye, leading, again, to a rotation toward the light. This physiological understanding of the biological mechanisms underlying phototaxis in *P. dumerilii* larvae is necessary for the insight that what appears as an external stimulus *eliciting* an until-then-inactive behavior is organized rather in the reverse fashion: these animals are constantly exploring their environment in the search for light, constantly changing directions. The external stimulus is then perceived as feedback from this exploratory, spontaneous behavior. In stark contrast of the implied activation of a response, the stimulus then only serves to *eliminate* a portion of the ongoing behavioral repertoire.

Thus, it appears as if behavior in a model for the Urbilaterian is organized in a fashion antithetical to the stimulus–response organization often assumed for nervous systems in general. These animals are first generating ongoing, random(-like), exploratory behavior that is modulated by subsequent reafferent feedback. This discovery may constitute the simplest, maybe even the earliest instantiation of an operant organization of behavior: generating a spontaneous action first and then evaluating its outcomes. If that were the case, "responses," if they actually exist, may only be rare and highly specialized, evolutionarily relatively late adaptations. In this view, the general concept of a stimulus–response organization of behavior is largely due to a combination of selection bias in which animal models and experiments are chosen for study and an inevitably superficial observation of the behavior.

Olfactory Reversal in Nematodes

If a stimulus–response organization of behavior were a particular evolutionary adaptation, e.g., evolved to speed up action selection in predictable situations, perhaps one can find examples of them in an animal model where we have an indication that evolution may have streamlined the pathways from stimuli to responses. One of the most well-studied genetic model organisms and so far the only adult animal with a complete connectome of its nervous system is the nematode worm *Caenorhabditis elegans* with its 302 neurons. The *C. elegans* connectome is dominated by feedforward connections from sensory neurons to motor neurons (Qian et al., 2011), so maybe the nervous system of this nematode is a promising candidate to find "responses" in the literal meaning of the word.

One well-characterized behavior in this nematode is reversal behavior. It occurs whenever the animal encounters aversive stimuli, such as certain odors. The circuit controlling this behavior can be described with just 4 neurons, their 44 chemical connections, and their electrical synapses. A central component of the system is a neuron called AVA. When AVA is active, the animal reverses its course. Sensory input to this neuron is provided by an olfactory neuron, AWC. For instance, if AWC is stimulated by an attractive odorant, it stops firing, such that AVA loses excitatory input and also stops firing, making reversals less likely. Conversely, activating AWC either experimentally or with an aversive odor increases the probability of reversals by synaptically activating AVA (Gordus et al., 2015). Two additional neurons are involved in this circuit, AIB and RIM, and the characterization of their role in the circuit is crucial for understanding the organization of olfactory-mediated reversal behavior in *C. elegans*.

The first interesting observation from the circuit connectivity is that there are more connections from the sensory AWC neuron to the AIB interneuron than to the reversal neuron AVA. This is unexpected, if the main function of nervous systems were to relay sensory information to motor centers. Imaging this circuit in immobilized worms in the absence of any stimuli reveals a complex pattern of correlated activity in all neurons. Interestingly, the neurons exhibit a sort of binary activity state, that for the most part is either on (neuron is active) or off (neuron is inactive). Quantifying the activity fluctuations in this circuit, one finds that there are three main states (of the eight theoretically possible) the circuit is commonly found in: just over 60% of the time the system is in "all on," roughly 20% is "all off," and for the remaining 20% it is in the state "only AIB on." This observation yields two insights: First, even without any stimulation at all, these network dynamics can generate spontaneous reversals without requiring any sensory input. Second, each olfactory stimulus reaching AWC will interact with the state the circuit currently happens to be in, rather than arriving in a quiescent circuit and triggering some neural activity that was not there before. The behavioral consequence of this interaction is not only the occurrence of spontaneous reversals, but also the occurrence of "spontaneous" nonreversals in

the presence of an aversive odor. In other words, the reversal circuit is probabilistic and without observing the nervous system of the worm, it is impossible to tell how spontaneous the observed behavior actually is.

Experimentally silencing either one or both of the interneurons in this circuit reveals that the role of AIB and RIM is to increase the variability of the reversal circuit. While the input to the circuit from the olfactory neuron AWC is always very precise and predictable if, e.g., an odor is presented, the activity of the reversal circuit always varies significantly and this variability is reduced if AIB, RIM, or both are silenced (Gordus et al., 2015). This discovery makes an excellent case for RIM and AIB being incorporated into the reversal circuit specifically to inject much needed variability into an otherwise maladaptively deterministic reversal circuit. Surprisingly, even though the feedforward connections dominate the connectivity also in this little circuit, the variability provided by the feedback connections dominates an adaptive feature of the behavior, its variability. This work adds *C. elegans* to the elongating list of animals, whose nervous systems are organized such that ongoing activity is merely modulated by external stimuli. In the nematode case, it appears that out of the four neurons comprised in this circuit, two exist for the sole reason to mitigate the effects that stimuli have on the behavior of the animal, to make the animal more autonomous with regard to its environment. If an animal with only 302 neurons, which, as in all other animals, make up the most energetically costly tissue, devotes 50% of a circuit to counter the effects of stimulus-response connections in its nervous system, then the implications of this discovery for the organization of behavior in animals generally cannot be underestimated.

Thus, contrary to the idea that a connectome dominated by feedforward connections from sensory to motor areas implies that it mainly computes motor output from sensory input, also the nervous system of *C. elegans* is best characterized by constantly changing, ongoing activity, much like many other nervous systems previously studied in this regard. It seems that even a numerically small feedback component provides a fundamental contribution to the overall architecture even of such feedforward-dominated networks. What does this mean for brains, such as those of mammals, whose neuroanatomy appears to be dominated by feedback loops?

Escape Responses

Few behaviors are as obviously under evolutionary selection pressures as escape responses: the antipredator behavior of prey in the presence of predators. Perhaps not surprisingly, this class of highly refined behaviors is particularly well studied because of their reproducibility in the lab, be it the C-start response in fish, mediated by the largest vertebrate neuron, the Mauthner cell (Korn and Faber, 2005; Schuster, 2012), the squid escape response mediated by their giant fiber system (Young, 1938), the crayfish giant fiber system that allows the animal to quickly propel itself out of harm's way by flipping its tail (Herberholz and Marquart, 2012), or the fly escape circuit also mediated by giant fibers originating in the optic lobes and making direct connections with the motor neurons that lead to both the raising of the wings and to the jump response of the legs (Hammond and O'Shea, 2007a; Card and Dickinson, 2008b, 2008a). All of these behaviors are easily observed in the lab, are mediated by conspicuous and easy to manipulate giant neurons and have, over the decades, shaped the way neuroscientists all over the world once conceptualized how behavior is organized: eliciting stimulus in, followed by behavioral response out (e.g., "brain function is ultimately best understood in terms of input/output transformations") (Mauk, 2000). However, studies in behavioral ecology and ethology have revealed that these famous behaviors are liable to exploitation (see, e.g., Catania, 2009, 2010; Jabłoński and Strausfeld, 2000, 2001), and/or are often more complex and flexible than initially thought (Card and Dickinson, 2008a, 2008b; Hammond and O'Shea, 2007a, 2007b; Herberholz and Marquart, 2012; Fotowat and Gabbiani, 2007). Searching for more escape behaviors under natural conditions to compare them with the better known laboratory examples, it was discovered that many escape behaviors contain elements of variability to make the escape trajectory less predictable for the predator (e.g., Royan et al., 2010; Domenici et al., 2008; Bateman and Fleming, 2014; Driver and Humphries, 1988; Humphries and Driver, 1970; Guerin and Neil, 2015; Highcock and Carter, 2014). This work demonstrates the adaptive value of such "protean" escape strategies and suggests that unpredictable prey is not only more difficult to catch, but also capable of injecting additional unpredictability into their behavior in the presence of predators. It thus appears as if the technically advantageous property of being highly reproducible not only renders an escape behavior liable to exploitation in the wild, but also has introduced a bias in neuroscience and psychology: reproducible laboratory behaviors are rarely representative of behaviors more generally, which tend to be much more variable and contain a larger degree of unpredictable spontaneity.

In the light of such data, it is tempting to postulate that the stimulus–response concept of animal behavior is little more than a laboratory artifact introduced into those sections of psychology and neuroscience that have isolated themselves from nonlaboratory behavior. Instead, converging data from multiple fields, model and nonmodel organisms in the laboratory and the wild suggest that at least all bilaterians have evolved the capability to inject a controlled amount of variability into their behavior to accomplish adaptive behavioral choice under a variety of situations. The influence of this "random number generator" ranges from changing minute aspects in the temporal dynamics of very simple behaviors to drawing from a set of different behaviors. The hallmark of such behavioral variability, the feature that confers its adaptive value is its independence from the environment, i.e., its spontaneity. For instance, the presence of a predator may increase a prey's behavioral variability, but not a specific escape trajectory or direction. In a novel environment, an animal may increase the variability of its foraging or exploratory behavior, but the environment cannot be used to predict the direction or duration of the foraging bout or the timing of the next exploratory behavior (see, e.g., Jacobs and Menzel, 2014 for a similar argument in a different research question).

The Interaction Between "Self" and "Nonself"

The observations discussed so far have led to a much more refined and sophisticated picture of how nervous systems organize behavior that can be circumscribed by the "emitted" and "elicited" dichotomy. From a neurobiological perspective, stimuli interact with and modulate ongoing activity in the nervous system, rather than triggering always the same cascades of neural activity in a more or less quiescent brain. This insight entails a number of important consequences.

For one, as behavior is based on neural activity, it is ongoing and continuous. The impression that behavior can be chunked into discrete acts (i.e., "units of behavior") is possibly due to both our tendency to categorize even continuous phenomena and the scale-free nature of neural activity, allowing us to find temporal patterns in behavior at different timescales.

Another consequence is that sensory information only needs to instruct the animal as to which kind of behaviors to favor over others. Which specific movements or actions are to be generated do not always have to be explicitly specified. Exploratory "trying out" with efficient feedback evaluation (i.e., operant behavior) is sufficient to quickly and reliably find the behavior that is best suited to solve the task at hand.

However, it is clear that some stimuli exert a more specific effect on the ongoing activity in the brain than others. There appears to be a continuum in the urgency or salience of events in an animal's environment, at one extreme end of which are stimuli so potent, that they appear to trigger a behavioral response, at least to the casual or superficial observer. Obviously, the coupling between a stimulus and an organism's behavior can vary from very loose to very tight both in terms of temporal coupling as in terms of which stimuli are followed by which behaviors, and vice versa. Most often, speed and efficiency are correlated with a tight environmental coupling, not only in stereotypical behaviors such as escape responses: the well-trained squirrel will much more quickly crack the nut than the naïve one. However, fast, efficient, and ultimately stereotypical behaviors are inflexible and liable to exploitation. Inasmuch as operant conditioning leads to the selection of efficient behavior and later to stereotypization of the behavior in a process called habit formation, operant processes are central in the animal's quest to trade-off efficiency for flexibility and unpredictability. This trade-off is essential for the survival and procreation of every organism and likely one of the most important ultimate causations behind adaptive behavioral choice.

Of course, there are ongoing processes inside the organism that can be ranked on a similar scale from hardly noticeable to directly influencing behavior. Hunger or circadian rhythms have at least as potent an effect on behavior as the presence of food, a predator or a potential mating partner. It is the complex interaction of processes generated by the animal itself (among which spontaneous behavior can be found) with processes generated by events outside of the animal that ultimately manifests itself as observable behavior. Importantly, the fact that we cannot tickle ourselves suggests that this distinction between self and nonself may be much more fundamental than one would at first assume. In their "reafference principle" von Holst and Mittelstädt famously proposed a mechanism by which animals use efference copies to distinguish between stimuli that are under their control (reafference) from those that are not (exafference) (von Holst and Mittelstaedt, 1950). The reafference principle is a fundamentally operant process by which animals not only distinguish between the stimuli they can control and those they cannot, but also the beginning of the operant process by which animals learn to bring novel stimuli under their behavioral control.

These considerations bring us back to another fundamental question Skinner formulated in the 1930s. Given the multitude of learning processes engaged during operant conditioning experiments, how many of them are similar to those taking place in classical conditioning (i.e., stimulus substitution) and how many, if any, are fundamentally different? Skinner himself lamented that he could not remove the lever his rats were pressing to eliminate one of the confounding factors (Skinner, 1935). Modern behavioral neurobiology in genetic model organisms has designed experiments specifically to address this problem and has discovered that the distinction between self and nonself is central to a purely operant learning mechanism.

Mechanisms of Plasticity in Operant Conditioning

Isolating the Components: Self-Learning

As Skinner noted, rats need a lever to press and thus they may learn about the food-predicting properties of the lever. Therefore, 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 environmental "contamination." One such preparation is tethered *Drosophila* at the torque meter as described earlier (Brembs, 2009b; Heisenberg et al., 2001; Wolf and Heisenberg, 1991; Heisenberg, 1994; Wolf et al., 1992; Heisenberg and Wolf, 1984; Wolf and Heisenberg, 1986). In the setup where the flies learn to control a punishing heat beam with their yaw torque in a homogeneous environment with no directional cues, the only contingency present is that between the behavior and the reinforcer/punisher. In this experiment, it is thus possible to isolate the operant contingency from all environmental contaminations. One may term such experiments "pure" operant conditioning experiments.

Using mutant, wild-type and transgenic animals, it was discovered that the canonical, cAMP-dependent synaptic plasticity pathway known from other learning experiments 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). Apparently, the common, evolutionary conserved mechanisms, discovered in classical conditioning experiments are not required for this pure operant learning. Instead, the animal has to rely on a different biochemical process if it is asked to learn about its own behavior, in the absence of any external

cues. With regard to the fundamental distinction between self and nonself discussed earlier, the content of the learning process is fundamentally about the animal's own behavior. Thus, one may call this process self-learning, i.e., the process by which value is assigned to a specific action or movement, such as the heat is assigning positive or negative value to left or right turning, respectively, in this experiment (Colomb and Brembs, 2010).

In the search for other components of the biological processes underlying self-learning, one may again refer to Skinner. This time his claim that language acquisition constituted a form of operant learning (Skinner, 1957): first exploratory, highly variable actions are being initiated (i.e., babbling) and then sensory feedback shapes the initiation of future behavior, reducing its variability (i.e., language). In what may appear at first to be a very superficial analogy, "mere homonyms, with at most a vague similarity of meaning" (Chomsky, 1959), operant self-learning in tethered flying *Drosophila* mimics these features of vocal learning: the animal first initiates highly variable, exploratory actions, then sensory feedback shapes the initiation of future behavior, reducing its variability. The Forkhead Box P2 (FOXP2) transcription factor is the first gene discovered to be involved in the development of speech and language (Fisher and Scharff, 2009; Lai et al., 2001). Importantly, the avian ortholog is also involved in song learning in birds (Scharff and Haesler, 2005; Haesler et al., 2007; Schulz et al., 2010), which has also been described as an operant behavior (Marler, 1991). The four different FoxP genes in vertebrates probably arose from serial duplications of a single ancestral FoxP gene after the separation from the invertebrate clades. The invertebrate FoxP ortholog corresponds most closely to the ancestral form of the gene at the base of the bilateria (Santos et al., 2011), thus lending itself to investigating the depth of the functional conservation among the members of the FoxP gene family. Mutant analysis and RNAi-mediated knockdown of the *Drosophila* ortholog, *dFoxP*, revealed its necessity specifically for self-learning (Mendoza et al., 2014), a phenocopy of the PKC manipulations described earlier. Thus, an excellent candidate for the ancestral function of FoxP genes, several of which are involved in acquiring the speech component of language in humans, is specifically involved in the isolated self-learning component of operant conditioning, but not in other forms of learning in *Drosophila*. These results are consistent with the hypothesis that the FoxP-dependent component of language evolved from an ancestral operant self-learning mechanism. The homology between invertebrate FoxP and its descendant genes, together with the similarities between habit formation and birdsong crystallization (Costa, 2011) prompts the postulation that habit formation in vertebrates may also be engaging this same mechanism.

Parallel developments to isolate the operant component have been made in the sea slug *Aplysia* (Nargeot et al., 1997, 2007, 2009; Nargeot, 2002; Nargeot and Simmers, 2010; Brembs et al., 2002; Lorenzetti et al., 2008, 2006; Nargeot et al., 1999c, 1999b, 1999a). As described earlier, reward signals were made contingent on spontaneous biting behavior, either in the intact animal or in isolated buccal ganglia. This procedure also excluded any external stimuli from contaminating this pure operant experiment, leading to self-learning. Such conditioning brings about two changes in the behavior of the animals/ganglia: both the total number of bites/BMPs is increased and the frequency of the rewarded behavior increases, relative to the other feeding behaviors. Electrophysiological studies discovered learning-related changes both in *When* and in *What* neurons. While the *What* neurons changed their excitability such that the behavior they promoted became more frequent (Brembs et al., 2002; Nargeot et al., 1999b, 1999a), the *When* neurons increased their electrical couplings as well as their excitability (Nargeot et al., 2009), such that not only the total number of behaviors increased, but they also became more stereotyped, losing the variability described earlier for the naïve animals in which the coupling between the *When* neurons was comparatively weak (Nargeot and Simmers, 2012). The self-learning mechanism in *Aplysia* hence entails direct modifications of the neurons controlling the type of behavior being generated as well as the variability of the behavior. In this way, we are starting to unravel the neuronal mechanisms behind selection by consequences (Skinner, 1981). These mechanisms appear to be evolutionary conserved as well, raising the possibility of a common, specifically operant learning mechanism for all bilaterians. Also in *Aplysia*, as in *Drosophila*, the canonical cAMP-dependent learning pathway discovered in classical conditioning is not involved, but PKC manipulations impair self-learning (Lorenzetti et al., 2008). Given that PKC is also involved in song-learning in birds (Yoshida et al., 2003; Sakaguchi and Yamaguchi, 1997), as is FoxP2, there is now strong evidence for such a conserved self-learning mechanism.

These results entail that there exists a dedicated biological mechanism that only occurs in neurons that are involved in actions and not in those processing the environment. With regard to the early questions of Skinner and his contemporaries, we can now say that we have learned that the different associative processes mediating the content of learning ("what is learned") in operant conditioning are mediated by different biological mechanisms: the molecular machinery involved in operant self-learning (i.e., PKC and FoxP to date) does not appear to be involved in any of the other types of learning studied so far. This insight was made possible by separating and isolating the self-learning process from any other processes that may take place during operant conditioning. Adding some of these components back to the experiment reveals interactions between these components that have a direct relation to the central efficiency/flexibility trade-off underlying adaptive behavioral choice discussed earlier.

Combining the Components: Composite Learning

Once the operant component has been isolated in "pure" operant learning experiments in which only self-learning can occur, it is comparatively easy to add a predictive stimulus and compare the resulting "composite" situation with the "pure" experiment. For instance, in tethered *Drosophila*, whenever the direction of turning maneuvers changes (e.g., from left to right turning attempts), 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 additional option to learn that one of the colors signals heat, and not only that its own behavior controls the heat. This situation is analogous to how rats in Skinner boxes may learn that the depressed lever signals food (and the undepressed lever no food) as well as that their lever pressing is predictive of the food reward. In contrast

to the "pure" experiment, this situation now requires the canonical cAMP cascade discovered in classical conditioning and is independent of any PKC signaling (Brembs and Plendl, 2008) or *dFoxP* function (Mendoza et al., 2014). Apparently, even though the experiment is just as operant as before the colors were added (in fact, if the flies were able to close their eyes to the colors, it would be the exact same, "pure" experiment), now seemingly "classical" learning mechanisms are engaged. This result suggests that as soon as learning of external stimuli becomes possible, these will be learned preferentially over behavioral cues, even in otherwise completely operant learning situations. As the content of these learning processes is fundamentally nonself, one may call these processes "world-learning," i.e., the process by which value is assigned to external stimuli, for instance, "green" being assigned an aversive value and blue an appetitive value in this experiment (Colomb and Brembs, 2010). In contrast to self-learning, world-learning can occur in both operant and classical learning experiments, depending on the operant control of the predictive stimuli (this will become a separate point of discussion later). Thus, the difference between operant and classical conditioning lies not only in the procedural differences between the two experiments but also in the content of the memory being formed: its content (i.e., "what is learned") decides which learning processes are engaged in the nervous system. How this content is learned, appears to be less relevant.

To ask whether the colors have been learned independently from the behavior with which they were learned, one can test the flies' preference of the unpunished color with an orthogonal behavior to that used during training. After all, to solve the "composite" situation, it is sufficient for the flies to learn that one of the colors is associated with the heat and then use whatever behavior necessary to avoid this color. Indeed, flies can avoid the punished color even with an orthogonal behavior. Conversely, when tested for a preference in turning direction (i.e., without colors) after composite conditioning, the flies do not reveal any preference (Brembs, 2009a). In the most Pavlovian sense, the flies seem to learn the color–heat contingency independently of the behavior with which it was acquired. They only learn about the world around them, without leaving much of an indication that the behavioral decision-making circuitry itself has been significantly altered, even though, of course, the entire situation is still just as operant as without the stimuli (Brembs, 2009a; Brembs and Heisenberg, 2000). Apparently, there is an inhibitory interaction between the world- and the self-learning processes during operant composite conditioning, such that only the effects of the world-learning process can be detected after operant composite conditioning. For this inhibition to occur, it is not sufficient that the colors are merely present: flickering the colors unrelated to the heat does not inhibit self-learning and mutants, which cannot learn the colors (but are not impaired in self-learning) do not show any sign of self-learning inhibition, even if the colors are predictive of the heat (unpublished observation). Thus, world-learning must be actively engaged, to inhibit self-learning.

To ask whether this inhibition is absolute or depends on other factors, such as time, one can extend the training to twice the duration used in traditional operant conditioning experiments in *Drosophila*. Testing for world- and self-learning effects after such extended training reveals that the inhibition of self-learning by world-learning is time-dependent: after extended training, the flies prefer generating the previously unpunished turning direction, even in the absence of the colors. In what appears to be an analogy to vertebrate "habit interference," the flies can no longer express the preference for the previously punished color with an orthogonal behavior (Brembs, 2009a). Interestingly, *FoxP* mutant flies are impaired in this type of habit formation, indicating that the self-learning that is taking place in pure operant conditioning is indeed the very same self-learning process that is inhibited in composite conditioning. One may interpret these results such 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. 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 (Brembs, 2009a).

Similar relationships have been observed in experiments with other animals. 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) (Packard and McGaugh, 1996; Hicks, 1964; Ritchie et al., 1950; Tolman et al., 1947, 1946). 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 (Berniker and Kording, 2008). There is a third field in which analogous results have been obtained. In experiments with rodents in operant chambers, extended training abolishes sensitivity to reinforce devaluation by the process of habit formation, which transforms goal-directed actions to habitual responses (Yin and Knowlton, 2006).

Besides the inhibitory interaction between world- and self-learning, there is a second interaction that can only be observed in composite conditioning. These experiments take advantage of the fact the world learning can occur in operant conditioning experiments (i.e., if they are composite experiments) and always occurs in classical conditioning experiments. If fly learning in composite situations is compared to situations where the same stimulus was trained classically, it is routinely observed that the stimuli are learned faster and to a higher level in the composite, than in the classical situation, even if the sensory input during training was identical between groups (Brembs and Heisenberg, 2000; Brembs and Wiener, 2006). This observation is reminiscent of 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 (Thorndike, 1898; Slamecka and Graf, 1978; Kornell and Terrace, 2007; James, 1890; Baden-Powell, 1908). The mechanism by which this facilitation of world learning occurs in composite conditioning compared to otherwise identical classical conditioning 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 the "generation effect."

It is straightforward to hypothesize that the difference in world-learning rate between classical and operant composite conditioning is due to the stimuli being presented exafferently in classical conditioning and reafferently in operant situations. Efference copies are generally proposed as one mechanism animals use to distinguish between exafferent and reafferent stimuli. If efference

copies are also used here to accomplish the generation effect in composite conditioning, it would entail that efference copies not necessarily always reduce the salience of reafferent stimuli as originally proposed (von Holst and Mittelstaedt, 1950), but can also enhance it, under certain circumstances, to accomplish a facilitation of world-learning. From these considerations it is a small step to postulate that when the relationship between a reafferent stimulus and the animal itself is concerned (such as in reafferent self-motion signals or self-tickling attempts) then efference copies serve to reduce the salience of the stimuli. However, when the relationship among different reafferent stimuli is concerned, then efference copies serve to increase their salience compared to the same stimuli presented exafferently.

Conclusions

In contrast to the long-held assumption that the organization of behavior is best characterized as the perception of a sensory stimulus followed by appropriate response (i.e., "sensorimotor hypothesis"), recent converging evidence from multiple systems and fields of study instead suggests that both ancestral and extant general brain function is best described in operant terms. Rather than specifying precise behaviors, sensory information—if at all present—interacts with ongoing neural activity to instruct the organism which type of spontaneous, exploratory behavior to generate. Evaluating the ensuing reafferent feedback modifies the nervous system such that ongoing neural activity patterns become biased toward activity that has generated increased appetitive and decreased aversive feedback in the past. The neurobiological mechanisms underlying both the exploratory, spontaneous behaviors and those underlying the modifications caused by the feedback are becoming increasingly understood, even on a molecular level. It is straightforward to hypothesize that the constant interaction between ongoing neural activity and the incoming sensory stream allows the organism to balance behavioral flexibility with efficiency to accomplish adaptive behavioral choice in an often unpredictably changing environment. The centrality of this trade-off is reflected in various fields discussing different aspects of this task. Evolutionary biologists study factors specifying whether flexible generalists or efficient specialists will prevail, psychologists search for transitions between flexible goal-directed actions and efficient habits, ecologists study at what point animals cease to efficiently exploit a resource to invest in costly exploration, computational neuroscientists unravel interactions between flexible model-based and efficient model-free learning processes. Animals and humans have evolved over millions of years to become experts in mastering this fundamental trade-off. Focusing on only one aspect of it fails to capture essential processes controlling behavior.

References

Baden-Powell, R., 1908. Scouting for Boys. C. Arthur Pearson Ltd, London.

Bateman, P.W., Fleming, P.A., 2014. Switching to plan B: changes in the escape tactics of two grasshopper species (Acrididae: Orthoptera) in response to repeated predatory approaches. Behav. Ecol. Sociobiol. 68, 457–465.

Bedecarats, A., Castro, J., Lade, Q., Cattaert, D., Simmers, J., Nargeot, R. 2015. No title. In: Society for Neuroscience Abstract (Chicago, IL), pp. 420.27.

Belanger, J.H., Willis, M.A., 1996. Adaptive control of odor-guided locomotion: behavioral flexibility as an antidote to environmental unpredictability. Adapt. Behav. 4, 217–253.

Benzer, S., 1967. Behavioral mutants of *Drosophila* isolated by countercurrent distriction. Proc. Natl. Acad. Sci. U.S.A. 58, 1112–1119.

Berniker, M., Kording, K., 2008. Estimating the sources of motor errors for adaptation and generalization. Nat. Neurosci. 11, 1454–1461.

Brembs, B., 1996. Chaos, cheating and cooperation: potential solutions to the Prisoner's Dilemma. Oikos 76, 14–24.

Brembs, B., 2009a. Mushroom bodies regulate habit formation in *Drosophila*. Curr. Biol. 19, 1351–1355.

Brembs, B., 2009b. The importance of being active. J. Neurogenet. 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 Biol. Sci.

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., Lorenzetti, F.D., Reyes, F.D., Baxter, D.A., Byrne, J.H., 2002. Operant reward learning in *Aplysia*: neuronal correlates and mechanisms. Science 296 (5573), 1706–1709.

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.

Card, G., Dickinson, M., 2008a. Performance trade-offs in the flight initiation of *Drosophila*. J. Exp. Biol. 211, 341–353.

Card, G., Dickinson, M.H., 2008b. Visually mediated motor planning in the escape response of *Drosophila*. Curr. Biol. 18, 1300–1307.

Carp, J.S., Tennissen, A.M., Chen, X.Y., Wolpaw, J.R., 2006. H-reflex operant conditioning in mice. J. Neurophysiol. 96, 1718–1727.

Catania, K.C., 2010. Born knowing: tentacled snakes innately predict future prey behavior. PLoS One 5, e10953.

Catania, K.C., 2009. Tentacled snakes turn C-starts to their advantage and predict future prey behavior. Proc. Natl. Acad. Sci. U.S.A. 106, 11183–11187.

Catania, K.C., 2008. Worm grunting, fiddling, and charming—humans unknowingly mimic a predator to Harvest Bait. PLoS One 3, e3472.

Chen, X.Y., Wolpaw, J.R., 1996. Reversal of H-reflex operant conditioning in the rat, 112, 58–62.

Chomsky, N., 1959. A review of B. F. Skinner's verbal behavior. Language (Baltim) 35, 26–58.

Colomb, J., Brembs, B., 2010. The biology of psychology: "Simple" conditioning? Commun. Integr. Biol. 3, 142–145.

Corcoran, A.J., Barber, J.R., Conner, W.E., 2009. Tiger moth Jams Bat sonar. Science 325 (5938), 325–327.

Costa, R.M., 2011. A selectionist account of de novo action learning. Curr. Opin. Neurobiol. 21, 579–586.

Dickinson, A., 1985. Actions and habits—the development of behavioral autonomy. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 308, 67–78.

Domenici, P., Booth, D., Blagburn, J.M., Bacon, J.P., 2008. Cockroaches keep predators guessing by using preferred escape trajectories. Curr. Biol. 18, 1792–1796.

Domjan, M., 2016. Elicited versus emitted behavior: time to abandon the distinction. J. Exp. Anal. Behav. 105, 231–245.

Driver, P.M., Humphries, N., 1988. Protean Behavior: The Biology of Unpredictability. Oxford University Press, Oxford, England.

Dunn, T.W., Mu, Y., Narayan, S., Randlett, O., Naumann, E.A., Yang, C.-T., Schier, A.F., Freeman, J., Engert, F., Ahrens, M.B., 2016. Brain-wide mapping of neural activity controlling zebrafish exploratory locomotion. Elife 5.

Fisher, S.E., Scharff, C., 2009. FOXP2 as a molecular window into speech and language. Trends Genet. 25, 166–177.

Fotowat, H., Gabbiani, F., 2007. Relationship between the phases of sensory and motor activity during a looming-evoked multistage escape behavior. J. Neurosci. 27, 10047–10059.

Glimcher, P., 2003. Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics. MIT, Cambridge, MA.

Glimcher, P.W., 2005. Indeterminacy in brain and behavior. Annu. Rev. Psychol. 56, 25–56.

Glimcher, P.W., Rustichini, A., 2004. Neuroeconomics: the consilience of brain and decision. Science 306 (5695), 447–452.

Gordus, A., Pokala, N., Levy, S., Flavell, S.W., Bargmann, C.I., 2015. Feedback from network states generates variability in a probabilistic olfactory circuit. Cell 161, 215–227.

Gorostiza, E.A., Colomb, J., Brembs, B., 2016. A decision underlies phototaxis in an insect. Roy. Soc. Open Biol. 6, 160229. http://dx.doi.org/10.1098/rsob.160229.

Grobstein, P., 1994. Variability in behavior and the nervous system. In: Ramachandran, V.S. (Ed.), The Encyclopedia of Human Behavior. Academic Press, New York, pp. 447–458.

Guerin, A.J., Neil, D.M., 2015. Escape trajectories of the rockpool prawn (Palaemon elegans) in response to visual and mechanosensory stimuli. Mar. Freshw. Behav. Physiol. 48, 145–161.

Haesler, S., Rochefort, C., Georgi, B., Licznerski, 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.

Hammond, S., O'Shea, M., 2007a. Escape flight initiation in the fly. J. Comp. Physiol. A. Neuroethol. Sens. Neural. Behav. Physiol. 193, 471–476.

Hammond, S., O'Shea, M., 2007b. Ontogeny of flight initiation in the fly Drosophila melanogaster: implications for the giant fibre system. J. Comp. Physiol. A. Neuroethol. Sens. Neural. Behav. Physiol. 193, 1125–1137.

Heisenberg, M., 1994. Voluntariness (Willkürfähigkeit) and the general organization of behavior. Life Sci. Res. Rep. 55, 147–156.

Heisenberg, M., Wolf, R., 1984. Vision in Drosophila. 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.

Herberholz, J., Marquart, G.D., 2012. Decision making and behavioral choice during predator avoidance. Front. Neurosci. 6, 125.

Hicks, L.H., 1964. Effects of overtraining on acquisition and reversal of place and response learning. Psychol. Rep. 15, 459–462.

Highcock, L., Carter, A.J., 2014. Intraindividual variability of boldness is repeatable across contexts in a wild lizard. PLoS One 9, e95179.

Hills, T.T., Kalff, C., Wiener, J.M., 2013. Adaptive Lévy processes and area-restricted search in human foraging. PLoS One 8, e60488.

von Holst, E., Mittelstaedt, H., 1950. Das Reafferenzprinzip. Wechselwirkungen zwischen Zentralnervensystem und Peripherie. Naturwissenschaften 37, 464–476.

Horn, C.C., Zhurov, Y., Orekhova, I.V., Proekt, A., Kupfermann, I., Weiss, K.R., Brezina, V., 2004. Cycle-to-cycle variability of neuromuscular activity in Aplysia feeding behavior. J. Neurophysiol. 92, 157–180.

Humphries, D.A., Driver, P.M., 1970. Protean defence by prey animals. Oecologia 5, 285–302.

Humphries, N.E., Sims, D.W., 2014. Optimal foraging strategies: Lévy walks balance searching and patch exploitation under a very broad range of conditions. J. Theor. Biol. 358, 179–193.

Jabłoński, P.G., Strausfeld, N.J., 2001. Exploitation of an ancient escape circuit by an avian predator: relationships between taxon-specific prey escape circuits and the sensitivity to visual cues from the predator. Brain Behav. Evol. 58, 218–240.

Jabłoński, P.G., Strausfeld, N.J., 2000. Exploitation of an ancient escape circuit by an avian predator: prey sensitivity to model predator display in the field. Brain. Behav. Evol. 56, 94–106.

Jacobs, L.F., Menzel, R., 2014. Navigation outside of the box: what the lab can learn from the field and what the field can learn from the lab. Mov. Ecol. 2, 3.

James, W., 1890. The Principles of Psychology. Holt, New York.

Jékely, G., Colombelli, J., Hausen, H., Guy, K., Stelzer, E., Nédélec, F., Arendt, D., 2008. Mechanism of phototaxis in marine zooplankton. Nature 456, 395–399.

Konorski, J., Miller, S., 1937a. Further remarks on two types of conditioned reflex. J. Gen. Psychol. 17, 405–407.

Konorski, J., Miller, S., 1937b. On two types of conditioned reflex. J. Gen. Psychol. 16, 264–272.

Korn, H., Faber, D.S., 2005. The Mauthner cell half a century later: a neurobiological model for decision-making? Neuron 47, 13–28.

Kornell, N., Terrace, H.S., 2007. The generation effect in monkeys. Psychol. Sci. 18, 682–685.

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.

Lorenzetti, F.D., Baxter, D.A., Byrne, J.H., 2008. Molecular mechanisms underlying a cellular analog of operant reward learning. Neuron 59, 815–828.

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.

Lum, C.S., Zhurov, Y., Cropper, E.C., Weiss, K.R., Brezina, V., 2005. Variability of swallowing performance in intact, freely feeding Aplysia. J. Neurophysiol. 94, 2427–2446.

Marler, P., 1991. Song-learning behavior: the interface with neuroethology. Trends Neurosci. 14, 199–206.

Martin, J.-R., Faure, P., Ernst, R., 2001. The power law distribution for walking-time intervals correlates with the ellipsoid-body in Drosophila. J. Neurogenet. 15, 205–219.

Mauk, M.D., 2000. The potential effectiveness of simulations versus phenomenological models. Nat. Neurosci. 3, 649–651.

Maye, A., Hsieh, C.-h., Sugihara, G., Brembs, B., 2007. Order in spontaneous behavior. PLoS One 2 (5), e443. http://dx.doi.org/10.1371/journal.pone.0000443.

McEwen, R.S., 1918. The reactions to light and to gravity in Drosophila and its mutants. J. Exp. Zool. 25, 49–106.

McNamara, J.M., Barta, Z., Houston, A.I., 2004. Variation in behaviour promotes cooperation in the Prisoner's Dilemma game. Nature 428, 745–748.

Mendoza, E., Colomb, J., Rybak, J., Pflüger, H.-J., Zars, T., Scharff, C., Brembs, B., 2014. Drosophila FoxP mutants are deficient in operant self-learning. PLoS One 9, e100648.

Miller, G.F., 1997. Protean primates: the evolution of adaptive unpredictability in competition and courtship. In: Whiten, A., Byrne, R.W. (Eds.), Machiavellian Intelligence II: Extensions and Evaluations. Cambridge University Press, Cambridge, MA, pp. 312–340.

Mitra, O., Callaham, M., Smith, M., Yack, J., 2009. Grunting for worms: seismic vibrations cause Diplocardia earthworms to emerge from the soil. Biol. Lett. 5, 16–19.

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. I. Contingent reinforcement modifies the functional dynamics of an identified neuron. J. Neurosci. 19, 2247–2260.

Nargeot, R., Baxter, D.A., Byrne, J.H., 1999b. 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., 1999c. Dopaminergic synapses mediate neuronal changes in an analogue of operant conditioning. J. Neurophysiol. 81, 1983–1987.

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., Petrissans, 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., 2012. Functional organization and adaptability of a decision-making network in Aplysia. Front. Neurosci. 6, 113.

Nargeot, R., Simmers, J., 2010. Neural mechanisms of operant conditioning and learning-induced behavioral plasticity in Aplysia. Cell. Mol. Life Sci.

Neuringer, A., 2004. Reinforced variability in animals and people: implications for adaptive action. Am. Psychol. 59, 891–906.

Packard, M.G., McGaugh, J.L., 1996. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. Neurobiol. Learn. Mem. 65, 65–72.

Qian, J., Hintze, A., Adami, C., 2011. Colored motifs reveal computational building blocks in the C. elegans brain. PLoS One 6.

Reynolds, A.M., Bartumeus, F., Kölzsch, A., van de Koppel, J., 2016. Signatures of chaos in animal search patterns. Sci. Rep. 6, 23492.

Ritchie, B.F., Aeschliman, B., Pierce, P., 1950. Studies in spatial learning. VIII. Place performance and the acquisition of place dispositions. J. Comp. Physiol. Psychol. 43, 73–85.

Royan, A., Muir, A.P., Downie, J.R., 2010. Variability in escape trajectory in the Trinidadian stream frog and two treefrogs at different life-history stages. Can. J. Zool. 88, 922–934.

Sakaguchi, H., Yamaguchi, A., 1997. Early song-deprivation affects the expression of protein kinase C in the song control nuclei of the zebra finch during a sensitive period of song learning. Neuroreport 8, 2645–2650.

Santos, M.E., Athanasiadis, A., Leitão, A.B., DuPasquier, L., Sucena, E., 2011. Alternative splicing and gene duplication in the evolution of the FoxP gene subfamily. Mol. Biol. Evol. 28, 237–247.

Scharff, C., Haesler, S., 2005. An evolutionary perspective on FoxP2: strictly for the birds? Curr. Opin. Neurobiol. 15, 694–703.

Schulz, S.B., Haesler, S., Scharff, C., Rochefort, C., 2010. Knockdown of FoxP2 alters spine density in Area X of the zebra finch. Genes Brain Behav. 9, 732–740.

Schuster, S., 2012. Fast-starts in hunting fish: decision-making in small networks of identified neurons. Curr. Opin. Neurobiol. 22, 279–284.

Shlesinger, M.F., 2009. Random searching. J. Phys. A Math. Theor. 42, 434001.

Shultz, S., Dunbar, R.I., 2006. Chimpanzee and felid diet composition is influenced by prey brain size. Biol. Lett. 2, 505–508.

Skinner, B.F., 1969. Contingencies of Reinforcement: A Theoretical Analysis. Appleton-Century-Crofts, New York.

Skinner, B.F., 1981. Selection by consequences. Science 213, 501–504.

Skinner, B.F., 1938. The Behavior of Organisms. Appleton, New York.

Skinner, B.F., 1935. Two types of conditioned reflex and a pseudo type. J. Gen. Psychol. 12, 66–77.

Skinner, B.F., 1937. Two types of conditioned reflex: a reply to Konorski and Miller. J. Gen. Psychol. 16, 272–279.

Skinner, B.F., 1957. Verbal Behavior. Copley Publishing Group.

Slamecka, N.J., Graf, P., 1978. Generation effect—delineation of a phenomenon. J. Exp. Psychol. Hum. Learn. Mem. 4, 592–604.

Thompson, A.K., Wolpaw, J.R., 2014a. Operant conditioning of spinal reflexes: from basic science to clinical therapy. Front. Integr. Neurosci. 8, 25.

Thompson, A.K., Wolpaw, J.R., 2015. Targeted neuroplasticity for rehabilitation. Prog. Brain Res. 218, 157–172.

Thompson, A.K., Wolpaw, J.R., 2014b. The simplest motor skill: mechanisms and applications of reflex operant conditioning. Exerc. Sport Sci. Rev. 42, 82–90.

Thorndike, E.L., 1898. Animal Intelligence. An Experimental Study of the Associative Processes in Animals. Macmillan, New York.

Tolman, E.C., Ritchie, B.F., Kalish, D., 1946. Studies in spatial learning. II. Place learning versus response learning. J. Exp. Psychol. 36, 221–229.

Tolman, E.C., Ritchie, B.F., Kalish, D., 1947. Studies in spatial learning. V. Response learning vs. place learning by the non-correction method. J. Exp. Psychol. 37, 285–292.

Wolf, R., Heisenberg, M., 1991. Basic organization of operant behavior as revealed in *Drosophila* flight orientation. J. Comp. Physiol. A 169, 699–705.

Wolf, R., Heisenberg, M., 1986. Visual orientation in motion-blind flies is an operant behavior. Nature 323, 154–156.

Wolf, R., Voss, A., Hein, S., Heisenberg, M., 1992. Can a fly ride a bicycle? Discussion on natural and artificial low-level seeing systems. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 337, 261–269.

Wolpaw, J.R., 2010. What can the spinal cord teach us about learning and memory? Neuroscientist 16, 532–549.

Wolpaw, J.R., Chen, X.Y., 2006. The cerebellum in maintenance of a motor skill: a hierarchy of brain and spinal cord plasticity underlies H-reflex conditioning. Learn. Mem. 13, 208–215.

Yin, H.H., Knowlton, B.J., 2006. The role of the basal ganglia in habit formation. Nat. Rev. Neurosci. 7, 464–476.

Yoshida, Y., Yamada, T., Sakaguchi, H., 2003. Activation of protein kinase C by the error signal from a basal ganglia-forebrain circuit in the zebra finch song control nuclei. Neuroreport 14, 645–649.

Young, J.Z., 1938. The functioning of the giant nerve fibres of the squid. J. Exp. Biol. 15, 170–185.