Operant Reward Learning in *Aplysia*

Björn Brembs[1]

Department of Neurobiology & Anatomy, University of Texas-Houston Medical School, Houston, Texas

Abstract

Anticipating the future has a decided evolutionary advantage, and researchers have found many evolutionarily conserved mechanisms by which humans and animals learn to predict future events. Researchers often study such learned behavior using conditioning experiments. The marine snail *Aplysia* has been at the forefront of research into the cellular and molecular mechanisms of classical conditioning. Recently, *Aplysia* has also gained a reputation as a valuable model system for operant reward learning. Its feeding behavior can be operantly conditioned in the intact animal as well as in reduced preparations of the nervous system. The reward signal relies on dopamine transmission and acts in conjunction with activity in an identified neuron (B51) to bring about operant memory.

Keywords

operant conditioning; *Aplysia*; reward; dopamine; learning; memory

As toddlers, we already know how to attract our parents' attention by pretending to cry. Learning to anticipate the consequences of our actions is central to shaping our personalities, and is accomplished through many different means, from processing social feedback to acquiring the motor skills for sports, crafts, or handiwork. In our daily lives, much of this fundamental type of predictive learning takes place unnoticed as the brain subconsciously processes the constant stream of stimuli, assesses the importance of each one, and cross-correlates them with our behavior. Some of the stimuli we encounter may have consequences that are more or less independent of our behavior: The smell of fresh coffee brewing in the morning, the sound of a dentist's drill in the waiting room, and dark clouds before a rainstorm are all signals of what is to come.

Obviously, we could not function without the capacity to learn the causes for future events. How does the brain accomplish this? How is the constant stream of relevant and irrelevant stimuli sorted and processed? To understand the neurobiological processes that perform these tasks, investigators must reduce the complexity of the environment to controlled, experimental circumstances, ideally involving only a single predictor and its consequence. Historically, such studies of predictive learning have been divided into two categories: studies in which the predictor is the subject's behavior (operant, or instrumental, conditioning) and studies in which it is a stimulus (classical, or Pavlovian, conditioning). In both cases, the predictor is repeatedly followed by its consequence and the subject learns that relationship.

EVOLUTIONARILY CONSERVED MECHANISMS

Decades of research into learning and memory have revealed that the capacity for predictive learning is so fundamental that even animals as distant from humans as worms, mollusks, and insects possess it. The presence of this evolutionarily conserved ability means that we might be able to understand certain basic human brain functions by studying these animals. As test subjects, they provide many technical advantages and allow experimentation that would be unthinkable with human subjects.

The marine snail *Aplysia* was introduced to neurobiology labs by Ladislav Tauc in the 1960s and was popularized as a research subject by Eric Kandel in the 1970s. The chief advantage of *Aplysia* as a research subject is that it has large neurons. Measuring up to 1 mm in diameter, they are easy to manipulate in a variety of ways. Conveniently, *Aplysia* also exhibits a surprising number of different learning capacities, including capacities for associative types of learning such as operant and classical conditioning. The results of studies on classical conditioning in *Aplysia* have proven so fruitful that Kandel was awarded the Nobel prize for physiology and medicine in 2000.

In classical conditioning studies, it was comparatively straightforward to trace the sensory pathways of the predictor (the conditioned stimulus) and its consequence (the unconditioned stimulus) into *Aplysia*'s nervous system. The learning had to take place in those neurons where the two stimuli converged (Antonov, Antonova, Kandel, & Hawkins, 2003; Walters & Byrne, 1983). In operant conditioning, the neural workings of the predictor (the operant behavior) are not so easy to trace.

APLYSIA FEEDING BEHAVIOR

Aplysia is a snail with virtually no natural predators. In its natural habitat, it is surrounded by its food (seaweed) and only has to raise its head and bite to eat. For these reasons, the animals exhibit only a

comparatively small repertoire of spontaneous behaviors that would be suitable for operant conditioning. The logical choice is to study feeding behavior. The situation for studying operant conditioning of *Aplysia* feeding behavior is almost ideal. First, when searching for food, the animals take random bites, even when no external stimuli trigger the bites (Kupfermann, 1974). Second, much of the neural network that generates the behavior (the central pattern generator, CPG) is known in great detail (Elliott & Susswein, 2002). It is situated in an aggregation of neurons (the buccal ganglia) located on the muscles that move *Aplysia*'s mouthparts (the buccal mass). Third, the sensory pathway of food stimuli involves the esophageal nerve (Schwarz & Susswein, 1986), which originates in the buccal ganglia; this morphology provides the potential for the necessary convergence of the behavior and the food reward in those ganglia. Finally, when removed from the animal, buccal ganglia continue to produce the neural patterns controlling the movements of the mouthparts (Morton & Chiel, 1993).

During experiments in which neural activity is measured when an intact animal is taking bites that fail to grasp food, the esophageal nerve shows little activity. However, when the animal grasps and swallows seaweed, bursts of electrical activity in the esophageal nerve accompany the ingestion of food (Brembs, Lorenzetti, Reyes, Baxter, & Byrne, 2002). Presumably, the esophageal nerve transmits information about the presence of food during swallowing to the buccal ganglia.

A VIRTUAL SEAWEED REWARD

The activity in the esophageal nerve that accompanies swallowing may be a reward signal. If so,

Aplysia that receive stimulation of the esophageal nerve immediately after each bite (contingent reinforcement), so that each stimulation might function as virtual food, should exhibit more biting behavior than a *yoked control* group, that is, a group in which the animals receive the same sequence of stimulation independently of their behavior. Indeed, in a study testing this prediction, this virtual food appeared to function as a reward for biting: Compared with both the yoked control group and a group that never received any stimulation, *Aplysia* that received the stimulation after each bite subsequently produced more bites in a test phase without any stimulation. This increase in biting was seen not only immediately after the training, but also 24 hr later (Brembs et al., 2002).

Apparently, the reward signal from the esophageal nerve converges on the neural activity in the buccal CPG responsible for the behavior. This finding simplified the task of investigating operant conditioning in *Aplysia*: Instead of behavioral experiments involving the entire animal, researchers could focus on a well-characterized network of comparatively large neurons, numbering in the hundreds. Consequentially, the next steps were to characterize the reward signal further and to find the neurons that are modified by the signal. Such detailed experiments required removal of the buccal ganglia from the animal so that researchers could study the neurons neurophysiologically and apply drug treatments that would not be feasible in the intact animal.

Isolated buccal ganglia in a petri dish (in vitro) containing artificial seawater continue to spontaneously produce, in seemingly random order, neural patterns of excitation (buccal motor programs, BMPs) that can be related to the different feeding-related movements

in the intact animal (Morton & Chiel, 1993). If these patterns are rewarded with the same type of electric stimulation of the esophageal nerve as in the experiment just described, in vitro operant conditioning takes place. Thus, isolated buccal ganglia that receive electrical stimulation after each BMP (contingent reinforcement) that resembles a bite in the intact animal (i.e., an ingestion-like BMP, or iBMP) produce more iBMPs than ganglia of the yoked control group (Nargeot, Baxter, & Byrne, 1997). This effect is blocked when a substance that blocks the effect of the neurotransmitter dopamine, methylergonovine, is added to the bath, implicating dopamine as the transmitter for the reward signal (Nargeot, Baxter, Patterson, & Byrne, 1999). Dopamine is also considered to be the prime transmitter for reward-related signals in humans and other mammals (Fiorillo, Tobler, & Schultz, 2003; O'Doherty, Dayan, Friston, Critchley, & Dolan, 2003).

CELLULAR MECHANISMS OF OPERANT REWARD LEARNING

Where in the feeding CPG in the buccal ganglion does dopamine act to make it produce more iBMPs? Neurons that can act as switches in the CPG, altering the output to produce different types of BMPs, are good candidates for playing a role in this function. Buccal neuron 51 (B51; Plummer & Kirk, 1990) is active late during an iBMP and is silent when the BMP resembles a movement that would reject an inedible item (a rejection-like BMP, or rBMP; Nargeot et al., 1997). Experimentally activating B51 during a BMP increases the likelihood that the BMP will become an iBMP. Conversely, silencing B51 during a BMP increases the likelihood that

the BMP will become an rBMP (Nargeot, Baxter, & Byrne, 1999a). Thus, B51 seems to be a pattern-switching neuron whose activation state largely determines the type of pattern the CPG will produce: If B51 is easily excited and likely to be active, iBMPs are more likely to occur, but if B51 is more difficult to activate, rBMPs are more likely to be produced.

After in vitro operant conditioning, B51 is more easily activated in ganglia that received contingent reward after iBMPs than in yoked controls (Nargeot, Baxter, & Byrne, 1999a). Thus, one mechanism by which in vitro contingent reinforcement may bring about operant learning is by modifying the properties of a pattern-switching neuron to render the CPG more likely to produce the rewarded behavior. Indeed, if stimulations of the esophageal nerve are made contingent simply upon activity in B51 (i.e., when this activity is experimentally induced and not part of a spontaneous BMP), the resulting increase in excitability in B51 alone is sufficient to reproduce some of the results of the in vitro operant conditioning just described (Nargeot, Baxter, & Byrne, 1999b). It is unknown how B51 changes if rBMPs are rewarded.

Is B51 relevant only in the isolated buccal ganglia, or does the in vitro preparation actually provide an accurate picture of the processes that occur inside the intact animal's central nervous system (i.e., in vivo)? B51 neurons from animals that have undergone the in vivo operant conditioning procedure show a higher excitability than B51 neurons dissected from yoked control animals (Brembs et al., 2002), mirroring the differences seen after in vitro operant conditioning. These experiments show that in vivo and in vitro operant conditioning of *Aplysia* feeding behavior produce the same kind of neural correlates of the operant memory.

Thus, we really can learn about the neural mechanisms of operant conditioning in vivo by studying parts of the isolated nervous system.

SINGLE-CELL OPERANT REWARD LEARNING

Studies of operant conditioning in *Aplysia* have covered all levels of complexity, from behavior, neural network, and single cells down to the molecules involved in changing the neurons' properties. *Aplysia*'s neurons are so big and robust that they can be taken out of the ganglion and cultured in petri dishes for several days. Based on the evidence for the convergence of a dopamine signal onto B51 activity during iBMPs, a single-cell analogue of operant conditioning can be established (Brembs et al., 2002), as in the following example. B51 is active late during an iBMP, and such activity can be triggered in cultured B51 neurons. Immediately following this activity, a pulse of dopamine is applied, to mimic the dopaminergic reward signal that follows an iBMP (in vitro) or a bite (in vivo) in the kind of experiments I described earlier. B51 neurons that have received seven such contingent dopamine applications show a higher excitability than B51 neurons that have received the dopamine exactly between two activations (Brembs et al., 2002). In other words, the effects of the contingent dopamine treatments parallel the effects found after both in vivo and in vitro operant conditioning. The molecular processes inside B51 that are involved in establishing these effects are currently under investigation.

Together, the results obtained thus far are consistent with the following model: In the intact animal, the dopamine-mediated food reward is contingent on B51 activity late during the rewarded behavior. The convergence of behavioral predictor and rewarding consequence in B51 leads to a modification of the biophysical properties of the neuron so that it is more likely to be active. These changes last for at least 24 hr. At least in part, these biophysical changes in B51, in turn, contribute to the increased frequency of bites seen after in vivo training.

THE FUTURE

The extent to which reward processing in *Aplysia* parallels reward processing in humans remains to be seen. As was the case with classical conditioning, the comprehensive, multilevel approach used to study operant conditioning in *Aplysia* has yielded some surprising parallels with operant conditioning in vertebrates. *Aplysia* offers potential for improved understanding of simple learning mechanisms that is unrivaled even in vertebrate research. At this time, *Aplysia* is the only system in which a convergence point of operant behavior and reward has been identified in the nervous system. We know a great deal about the neural network that gives rise to *Aplysia*'s biting behavior and the mechanisms through which the reward acts to modify components in this neural network to generate operant learning. Consequently, the deduced model of this system is unparalleled in quality and in the predictions it makes.

However, much remains to be learned about operant reward learning of *Aplysia* feeding behavior. Surely, B51 cannot be the only site of plasticity (i.e., change induced by operant learning) in the buccal ganglia. If it is not, what is its quantitative contribution to the total learning process? Where are the other sites of plasticity? Will the mechanisms in the other sites be

similar to or very different from those in B51? How many sites are there, and at what stage in the generation of behavior are they involved? Where are the sites of interaction with classical conditioning, if there are any?

Studying operant reward learning in *Aplysia* may be especially useful for another reason: In humans, reward learning can also lead to the development of maladaptive behavior patterns, such as addiction. Investigating the critical mechanisms underlying reward learning has been a prominent theme of psychological and neuroscience research over the past century. It is known that the dopaminergic system is crucial for the development of most types of addiction, but the subcellular processes involved are still largely unknown. Using *Aplysia* as a model system for studying this problem might be fruitful.

Recommended Reading

Brembs, B., Lorenzetti, F.D., Reyes, F.D., Baxter, D.A., & Byrne, J.H. (2002). (See References)
Nargeot, R., Baxter, D.A., & Byrne, J.H. (1997). (See References)
Nargeot, R., Baxter, D.A., & Byrne, J.H. (1999a). (See References)
Nargeot, R., Baxter, D.A., & Byrne, J.H. (1999b). (See References)
Nargeot, R., Baxter, D.A., Patterson, G.W., & Byrne, J.H. (1999). (See References)

Note

1. Address correspondence to Björn Brembs, Department of Neurobiology & Anatomy, University of Texas-Houston Medical School, Houston, TX 77030; e-mail: bjoern@brembs.net.

References

Antonov, I., Antonova, I., Kandel, E.R., & Hawkins, R.D. (2003). Activity-dependent presynaptic facilitation and hebbian LTP are both required and interact during classical conditioning in *Aplysia*. *Neuron, 37*, 135–147.

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.

Elliott, C.J., & Susswein, A.J. (2002). Comparative neuroethology of feeding control in molluscs. *Journal of Experimental Biology, 205*, 877–896.

Fiorillo, C.D., Tobler, P.N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science, 299*, 1898–1902.

Kupfermann, I. (1974). Feeding behavior in *Aplysia*: A simple system for the study of motivation. *Behavioral Biology, 10*(1), 1–26.

Morton, D.W., & Chiel, H.J. (1993). The timing of activity in motor neurons that produce radula movements distinguishes ingestion from rejection in *Aplysia*. *Journal of Comparative Physiology A, 173*, 519–536.

Nargeot, R., Baxter, D.A., & Byrne, J.H. (1997). Contingent-dependent enhancement of rhythmic motor patterns: An in vitro analog of operant conditioning. *Journal of Neuroscience, 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. *Journal of Neuroscience, 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. *Journal of Neuroscience, 19*, 2261–2272.

Nargeot, R., Baxter, D.A., Patterson, G.W., & Byrne, J.H. (1999). Dopaminergic synapses mediate neuronal changes in an analogue of operant conditioning. *Journal of Neurophysiology, 81*, 1983–1987.

O'Doherty, J., Dayan, P., Friston, K., Critchley, H., & Dolan, R. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron, 38*, 329–337.

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. *Journal of Neurophysiology, 63*, 539–558.

Schwarz, M., & Susswein, A.J. (1986). Identification of the neural pathway for reinforcement of feeding when Aplysia learn that food is inedible. *Journal of Neuroscience, 6*, 1528–1536.

Walters, E.T., & Byrne, J.H. (1983). Associative conditioning of single sensory neurons suggests a cellular mechanism for learning. *Science, 219*, 405–408.