



## Is Positive Reinforcement Pseudoscience?

Gunnar Newquist\*

Brain2Bot, Inc., 100 N. Arlington Ave. Ste. 350, Reno, NV 89501, USA

**\*Corresponding Author:** Gunnar Newquist, Brain2Bot, Inc., 100 N. Arlington Ave. Ste. 350, Reno, NV 89501, USA, **Email:** [g.newquist@brain2bot.com](mailto:g.newquist@brain2bot.com)

**Abstract:** Reinforcement forms the basis for our current principles of learning. However, the definition of reinforcement is circular logic. As such, reinforcement bears no explanatory power, and therefore, cannot be used as a scientific principle of neuroscience. If we wish to understand the brain, we must replace the principle of reinforcement with a falsifiable principle that is consistent with results from controlled experiments.

**Keywords:** Reinforcement, Reward, B.F. Skinner, Pavlovian (Classical) Conditioning, Operant Conditioning, Pseudoscience

### ARTICLE

One of the hallmarks of modern science is falsifiability. In science, there must be a clear-cut way to determine if you are wrong all the way from a preliminary hypothesis up to a grand unifying theory. For example, in medicine, one simply compares a treatment in question to a placebo control. If the treatment performs better than a placebo (under unbiased conditions), it must be doing its purported job. If not, the treatment and hypothesis concerning mechanism of effectiveness are wrong. Try something else.

But this fundamental principle of falsifiability is being overlooked in neuroscience. Part of the problem cuts to the very foundation of our field, and it relates to the most basic processes of how a brain creates behavior.

B. F. Skinner proposed a theoretical process he called “reinforcement”, which he claimed was the process that controls conditioning and, ultimately, behavior [1]. In Skinner’s explanation of conditioning, a reinforcing stimulus that follows a behavior makes it more likely that that preceding behavior will occur again in the future. Here is a summary of the method he used to demonstrate reinforcement, called Operant or Instrumental conditioning:

Wait for the subject to perform some desired response. After the desired response occurs, reward the subject with some commodity, such as food. After repeatedly rewarding the response, positive reinforcement occurred if the

frequency of rewarded response goes up. Hence, learning was successful.

One of the canonical demonstrations of this reinforcing effect uses a rat and a lever. During training, food is delivered only after the rat presses a lever. After training, the frequency of lever pressing goes up, and therefore, positive reinforcement is concluded.

These results are powerful, and the explanation is intuitive. The rat was not pressing the lever much, if at all, before food delivery. Then, after a few rewarded lever presses, *Presto!* The rat presses the lever over and over again. The food came after the lever press, so the rat must have associated the lever press behavior with the reward. The experimenter concludes that reinforcement occurred.

Well, not so fast. . .

Another method for assessing conditioning already existed before the birth of Skinner’s Operant conditioning. That method was Pavlovian or Classical conditioning. In Pavlovian conditioning, there is no response requirement. Two stimuli just have to be presented, paired in time. After training, one of those stimuli is presented alone. If a change of behavior occurs compared to the naïve state, learning was successful. (Of course, this change is compared to a control stimulus that is presented unpaired.)

For the rat/lever example, the Pavlovian version is to present the lever and the food together,

without regard to whether the rat actually pressed the lever during training. Now bear in mind, the rat can be doing anything it is that rats do at the time of food delivery, and they do not need to do anything at all to receive food. What is the result of this free food delivery method? The rat presses the lever over and over, just as it would in the Operant conditioning task!

Now, this result bares reflection. The rat did not press the lever before food appeared. In some Pavlovian experiments, the lever is not even available before food delivery. In Pavlovian conditioning, the food could have rewarded the rat for anything else it was doing: scratching, sniffing, staring blankly at the wall. But the food did not reward other behaviors. The rat pressed the lever. The rat seemed compelled to press the lever without ever being rewarded for pressing levers. Positive reinforcement cannot explain Pavlovian conditioning.

Pavlovian conditioning in learning is much like the placebo effect in medicine: it is the simplest explanation for a change in behavior. A scientist must rule out Pavlovian conditioning effects with a Pavlovian control group in a conditioning test before he or she can claim any additional effects, just like a scientist must rule out placebo effects with a placebo control in a drug test before he or she can say a specific treatment had any additional effects. Medicine without a placebo control is not medicine. It is pseudoscience.

Skinner claimed that his Operant conditioning brought behavior into the realm of scientific inquiry [2]. In an enormous irony, Skinner never published a single control group comparing his Operant conditioning to Pavlovian conditioning to support his “scientific” reinforcement claim. Over 30 years after Operant conditioning was invented, other researchers finally began publishing the Pavlovian method for lever pressing, which they called autoshaping [3]. Autoshaping clearly demonstrated that reinforcement was not necessary to explain the development of the canonical lever press response. Today, however, now nearly 80 years since Skinner’s first Operant conditioning publications, positive reinforcement continues to be used to describe how a rat learns to press a lever [4-9].

**If autoshaping was so clearly Pavlovian conditioning, why did Skinner’s reinforcement remain afterward?**

Though autoshaping demonstrated that positive reinforcement was not necessary to explain

lever pressing *per se*, autoshaping did not eliminate positive reinforcement from the list of possible explanations for conditioning. It couldn’t. And neither could any other set of experiments. That is because the very definition of reinforcement is unfalsifiable:

**Reinforcement:** *A reinforcing stimulus following a behavior makes it more likely that the behavior will occur again in the future.*

If the stimulus was a reinforcer, it will increase the previous behavior. How does one determine if the stimulus was a reinforcer? It increased the behavior. What process increased behavior? Reinforcement. If the premises are true, the conclusion must be true. Reinforcement is a textbook example of circular logic. Circular logic is not falsifiable, and therefore, cannot be a basic principle of behavior.

But now there is a problem. Skinner’s unfalsifiable, unscientific reinforcement theory *is* being used as a basic principle of behavioral science. Today, both economic utility theories and reinforcement theories fall into this circular logic trap by using reinforcement as the explanation for both behavior and the underlying learning. Economic utility theories state that behaviors chosen by an organism maximize a hypothetical numerical value, a value that is called “utility”, which is the same logical statement as maximizing reinforcement value. Reinforcement theories state that learning changes behavior to maximize future reward, which is the same logical statement as maximizing future reinforcement. Reinforcement, whether it be direct, as a utility value, or indirect, as an estimate of future reward value, is now assumed to capture all of the relevant factors influencing choice in the animal [10]. Under current theory, reinforcement is no longer a hypothesis to be tested with a control group, but part of the definition of learning.

But reinforcement is not a principle. Reinforcement is not science. The above theoretical frameworks do not provide explanatory power. If we are using an untestable theory to describe a fundamental process in the brain, that means we do not actually understand the brain. As captivating as reinforcement has been for the last 80 years, it has been hindering our progress in understanding ourselves. Applications as diverse as education to dog training and even cutting edge artificial intelligence software rely on neuroscience to

provide scientific learning principles from which to build innovations. Today, we need to critically evaluate whether or not we can falsify all of our most fundamental principles. The fate of behavioral neuroscience rests on our ability to produce a new generation of testable theories and actually rule them out. Neuroscience and all of its applied technology depends upon it.

### REFERENCES

- [1] Skinner B.F., *The behavior of organisms: An experimental analysis*. New York: Appleton-Century (1938).
- [2] Skinner B.F., *The experimental analysis of behavior*, *American Scientist*.45, 343-71 (1957).
- [3] B. Schwartz and E. Gamzu, Pavlovian control of operant behavior: An analysis of autoshaping and its implications for operant conditioning. In: *A Handbook of Operant Behavior*, Honig W.K., Staddon J.E.R., eds , Englewood Cliffs NJ: Prentice Hall, Pp76-108,(1977).
- [4] Costello M.R., Reynaga D.D., Mojica C.Y., Zaveri N.T., Belluzzi J.D., and Leslie F.M., Comparison of the Reinforcing Properties of Nicotine and Cigarette Smoke Extract in Rats, *Neuropsychopharmacology*. 39, 1843-51 (2014).
- [5] McNamarra A.A., Johnson L.E., Tate C., Chiang T., Byrne T., Acquisition of operant behavior in rats with delayed reinforcement: A retractable-lever procedure, *Behavioral Processes*. 111, 37-41 (2015).
- [6] Roberts D.S., Gabriele A., Zimmer B., Conflation of Cocaine Seeking and Cocaine Taking Responses in IV Self-administration Experiments in Rats: Methodological and Interpretational Considerations, *Neurosci. Biobehav. Rev.* 37, 2026-36 (2013).
- [7] S.T. Schepers and M.E. Bouton, Hunger as a Context: Food Seeking That Is Inhibited During Hunger Can Renew in the Context of Satiety, *Psychol. Sci.* 1-9 (2017).
- [8] Schulz D., Henn F.A., Petri D., and Huston J.P., Rats bred for helplessness exhibit positive reinforcement learning deficits which are not alleviated by an antidepressant dose of the MAO-B inhibitor deprenyl, *Neuroscience*. 329, 83-92 (2016).
- [9] Trezza V., Campolongo P., and Vanderschuren L.J., Evaluating the rewarding nature of social interactions in laboratory animals, *Dev. Cogn. Neurosci.* 4, 444-58 (2011).
- [10] Lee D., Seo H., and Jung M.W., Neural basis of reinforcement learning and decision making, *Annu. Rev. Neurosci.* 35, 287-308 (2012).

**Citation:** Gunnar Newquist. *Is Positive Reinforcement Pseudoscience?*. *ARC Journal of Neuroscience*. 2017; 2(3):4-6. doi:dx.doi.org/10.20431/2456-057X.0203002.

**Copyright:** © 2017 Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.