

## LEARNING PART2

Last time, we talked about classical conditioning. For example a rat salivates to cheese naturally, so cheese is the unconditioned stimulus, and the salivation is the unconditioned response. If we turn on a light every time we present the cheese, after a while the rat will start to salivate just when the light is turned on because the light announces the cheese. This is **classical conditioning**, and the light is a conditioned stimulus whereas the salivation is conditioned response. Classical conditioning is a passive process. The rat in the example does not do anything, and does not initiate any events. But a lot of what we do is active, meaning is initiated by us.

In order to account for those behaviors that are initiated by us, we will talk about a different view of learning, which is **operant conditioning**. In operant conditioning, the organism changes what happens. Operant conditioning is similar to classical conditioning but has very different assumptions about the learning process. Animals, like us, can initiate behavior; then the behavior has consequences, some of which we like, and some of which we don't like. If I initiate a behavior and it elicits a positive consequence/reward, I am going to redo this behavior. According to this view, we learn by experiencing the consequences of our behavior.

### Some important terms

**Operant:** behaviors we perform

**Reinforcers:** things that happen as a result of our behavior.

The general rule is that if a behavior creates something pleasant for us then do we it again.

### E. Thorndike's law of effect

This law summarizes this view of learning: if the response in the presence of a stimulus is followed by a satisfying event, the association between the stimulus and response is strengthened; if the response is followed by an annoying event, the association is weakened.

How did Thorndike study this? Thorndike created a **puzzle box** about which we watched a video clip. This is a complicated box, and typically Thorndike would put a cat in the complicated box. Of course, the cat wants to get out. So the cat would explore the box, walk around, and then at some point, it would pull the lever, which would result in the box opening and the cat would walk out to a food plate. Being able to get out of the box and to get the food is rewarding, so next time the cat is put in the box, it will know what to do to get out. A well practiced cat is able to get out of the box quickly. There is one very important point: the cat figures out how to get out by trial and error. It's not that the cat had some insight, some creativity, or some deep thought,

it just tried random things until something worked. This is important because it says that creativity and thinking are superfluous concepts for understanding how operant conditioning works.

### **BF Skinner and operant conditioning**

Thorndike came up with the law of effect, but operant conditioning is more often associated with BF Skinner. Skinner took the law of effect of Thorndike, named it operant conditioning, and began to study it systematically.

There are several differences between classical conditioning and operant conditioning.

- In classical conditioning, the stimulus precedes and elicits the response, whereas in operant conditioning the stimulus follows and strengthens the response : the animal behaves and gets something.
- In classical conditioning, learning is about associations whereas in operant conditioning, learning is about consequences of behavior
- Classical conditioning is associated with Pavlov, operant conditioning is associated with Skinner

Skinner's work on operant conditioning revolutionized psychology. Thorndike had the idea, but it only became powerful after Skinner's work. Skinner named it operant conditioning and created behaviorism, which dominated the field of psychology from 40s to 60s. Before behaviorism, psychology was dominated by psychoanalysis, which emphasized the role of the unconscious. For Skinner, the unconscious, thoughts, beliefs, ideas, all of that was irrelevant to the understanding of behavior. It's all about behavior, hence the name behaviorism.

The Skinner box was more sophisticated than Thorndike's puzzle box: it might have speakers that make noises, an electric grid that sends electric shocks, lights of different colors, a lever that dispenses food pellets when pressed. All these in the rat's environment provide feedback about behavior.

For Skinner and others, operant conditioning was sufficient to explain all of behavior, whether it was the behavior of humans or the behavior of rats. Some critics argued that not all of behavior was that simple, consisting of just rewards and punishments. In response, the behaviorists came up with more complex ideas to explain more complex behaviors, while maintaining that behaviorism could explain all of behavior. **Shaping** can get animals to do complicated things. It consists of providing reinforcement for successive approximations of the desired behavior. An example of shaping is potty training. A parent might reward a child by saying "good job" when the child goes to the bathroom, then when the child sits on the toilet, etc, until the child learns what to do.

## Reward structures

There is a variety of ways to provide reinforcement. Reinforcement can be **continuous**, meaning that reinforcement is given every single time the target behavior is performed. But in life, behavior is not rewarded or punished every time. **Partial reinforcement** is more common than continuous reinforcement. Continuous reinforcement is quickly learned, but it is more vulnerable to extinction: as soon as the reinforcement stops, the behavior will stop. Partial reinforcement is slowly learned; however, it is much less vulnerable to extinction.

Ratio vs. interval schedules: in a **ratio schedule**, the reinforcement is given after the behavior has been initiated a certain number of times. In an **interval schedule**, the reinforcement is given after a certain amount of time has passed, regardless of how much the behavior has been initiated.

Both ratio and interval schedules can be **fixed or variable**. For a fixed ratio schedule, the reinforcement might come say every 3 times the behavior is performed. For a variable ratio schedule, the reinforcement might come after a random number of times the behavior is performed: it might be after 2 times, then 10 times, then 4 times. For a fixed interval schedule, the reinforcement might come every 10 minutes, whereas in a variable interval, it might be 10 minutes, then 5 minutes, then 19 minutes.

### Some examples:

- **Fixed interval:** a child asks his parents every 60 seconds "are we there yet"
- **Variable interval:** a child asks his parents at random times "are we there yet"
- **Variable ratio:** gambling and lottery, and other things of that nature where you don't know when you will get reinforced, but you keep putting money in hoping that this time it will happen

In ratio schedules, the behavior is done more often than in interval schedules because the reward depends on how often the behavior is done. A rat can learn that it has to keep pushing until the 10<sup>th</sup> push for the food pellet to come, so it will keep pushing. Variable ratio schedules work in a similar way. In interval schedules, the behavior is not as rapid because the reinforcement depends on time. Imagine you were buying concert tickets, and that the tickets were on sale only between 8 and 9am. You would call a lot (behavior) between 8 and 9 am, if you couldn't get tickets that day, you would stop calling until the next day at 8 am. It would not make sense to keep calling after 9am.

## Reinforcement and punishment

There are different ways to think about reinforcement styles. There are different aspects of reinforcement and its flip side, punishment. **Reinforcement** increases behavior,

**punishment** decreases behavior. Both reinforcement and punishment can be **positive and negative**. Be careful: here positive does not mean good and negative bad; they mean something different in behaviorism. **Positive** means the stimulus happens, negative means that a stimulus is taken away. For example: if a rat is given cheese every time it presses a lever, the cheese is a positive reinforcement because it appears as a result of the behavior (pushing the lever). If a rat is given an electric shock (punishment) every time it pushes a lever, the electric shock is a positive punishment because the shocks happen as a result of the behavior. If the rat's cheese is taken away as a result of a lever press, it's a negative punishment: something the rat likes is taken away. If the electric shocks stop every time a rat presses a lever, it's a negative reinforcement; something bad is taken away as a result of the behavior.

In summary:

- the law of effect was an original idea of E. Thorndike
- Shaping means reinforcing successive approximations of the desired behavior
- Reinforcement can follow a fixed or variable schedule

### **Criticisms of behaviorism**

Radical behaviorists (Skinner and Watson, and their followers) believed that all behavior resulted from reinforcement. Watson went as far as saying that he can make a lawyer or a beggar or a doctor or a thief of any child just by controlling the child's reinforcements. Skinner wrote a novel describing his vision of an ideal society based on the principles of behaviorism. They assume that the mind is a blank slate, that everything someone becomes depends on the environment and that there is no learning without reinforcement. This was a quite extreme position, and some researchers began to challenge behaviorism.

**Tulman's experiments:** If a rat is put in a maze and there is cheese at the end of the maze, according to behaviorism, the rat will be much faster overtime at getting to the end of the maze because it has been rewarded for getting to the end. Rats that have not been rewarded are typically slower. Tulman wanted to challenge the idea that learning cannot happen without reinforcement. He had a group of rats run through a maze without giving them any cheese for 8 days. On day 9, he put cheese at the end of the maze. If indeed the rats had not learned anything because they had not been rewarded during the first 8 days, then they should be really slow. If on the other hand the rats have learned something, then they should be fast with finding the cheese, even though they have not been rewarded before. That's what Tulman found: on day 9, the rats were really fast at finding the cheese, so they had learned about the structure of the maze without reinforcement.

**Observational learning:** these experiments showed that behaviors can be acquired or modified just by observing others. Animals and children learn by observing. They don't

have to do the actual behavior to get reinforcement, they just see others do it, and they do it too. There are many things we learn that way. For instance monkeys raised in the lab aren't afraid of snakes but wild monkeys are. When a wild monkey is put in a cage with a lab monkey, and a snake is presented to them, the lab monkeys learn to be afraid because they have observed the wild monkey's fear reaction. **Bandura's experiment on observational learning:** this study is called the bobo doll. Kids watched a clip of an adult with a bobo doll (a plastic inflated doll) and there were 3 conditions; in one condition, the kids watched a clip of an adults who were aggressive with doll (hitting the doll, kicking it); in the second condition, the adults did gentle things with the doll (like hugging it), and the third condition was a control condition for which there was no doll. After watching the clips, the kids are sent to a different room where there is a bobo doll and many other toys; the researchers were interested in the number of aggressive acts towards the doll. The kids who watched the gentle tape showed less aggression towards the doll than kids in the control condition and kids who watched the aggressive tape. Kids who watched the tape in which the adults were aggressive behave very aggressively towards the doll: they not only reproduced the violent acts performed by the adults on the tape, but they also came up with new ways to be aggressive. These examples clearly illustrate that the view that only reinforcement can predicted behavior is limited. The fate of behaviorism provides an important lesson about progress in science. It's not that the findings of behaviorism were to be removed completely; instead they were to be nuanced and expended.