Module: Psychological Foundations of Mental Health

Week 1
Introduction to cognitive psychology

Topic 2
The heyday of behaviourism: operant learning – Part 3 of 3

Professor Richard Brown
Department of Psychology, King's College London

Lecture transcript

Slide 3
Let's get back to rats and Skinner's work. Like Thorndyke before him, Skinner was concerned with operant learning. Unlike Thorndyke, who was interested in animals' seeming ability to solve problems with his puzzle boxes, Skinner reduced the measurements of behaviour to simple parameters-- the speed at which a response was made and repeated, its intensity, and to a lesser extent, it's duration.

The puzzle box was refined to the operant chamber that is still used today in various forms and is commonly called a Skinner box. Skinner's preferred lab animals were the pigeon, that was well adapted to learn to peck at a light stimulus at a response, and a rat, which could press a lever with its paw. Operant chambers have subsequently been built that can exploit to the behavioural repertoire of just about every animal species.

The animal is put in the box only for the time of the experiment. Although designs vary, all boxes have the following common features. There is a means of providing a stimulus, often a siren or a light; a means of the animal making a response, such as a lever that can be pressed; and a means of delivering a reinforcer or punishment, often a small quantity of food delivered or a mild electric shock. Finally, the box is linked to a controller, now a computer, that determines the presentation of the stimuli, controls the delivery of the reward or punishment, and records the animal's response.

Slide 4
One of the legacies of the work of Skinner and other behavioural psychologists has been a detailed understanding of the precise relationship or contingency linking behaviour and the consequent reinforcer or punisher in terms of how and when it is delivered. This work includes the description of a large number of so-called schedules of reinforcement that appear to have universal applicability. They can both describe and predict the behaviour of rats pressing a lever and of human gamblers pouring their money into a slot machine.

We will look at a few examples here and some of the simpler schedules used and defined by Skinner. I will use the word reinforcement in general sense to describe any outcome which changes behaviour. The simpler schedule is one where reinforcement occurs every time a response is made, so-called continuous reinforcement schedule or CRF.

Transcripts by 3Playmedia

In such a schedule, the rat would receive a food reward every time it pressed the lever in response to the stimulus. In a human example, we get a cup of coffee dispense every time we put our money into a coffee machine, usually. More commonly, both in laboratory experiments and in the real world, reinforcement occurs only after some responses, not after all-- called intermittent schedules or partial reinforcement schedules.

Two main sets of parameters define the main types of schedules. The first parameter is based on the timing or quantity of responses, whether the reinforcer is given after a set number of responses, a ratio schedule, or after a set amount of time, an interval schedule. The second parameter is based on the certainty or predictability that the response will lead to reinforcement, whether the reinforcer is given after a fixed number of responses or amount of time, or after a variable number of responses or time.

These parameters can be combined in many different ways, but we will consider the four main ones here. In a fixed ratio or FR schedule, the rat gets a fruit reward after a fixed number of responses. In an FR2 schedule, the reinforcer is delivered every second response. In FR10, every 10th response, and so on. You can see that a continuous reinforcement schedule is, in fact, a fixed ratio or FR1 schedule, where the rat gets a reward every time it presses a lever.

In a variable ratio or VR schedule, the reinforcer is given, on average, every nth response, rather than precisely every nth response, as it is in a fixed ratio schedule. So for a VR10 schedule, the reinforcer may be given at the first response on some occasions and the 20th on others. However, with repeated trials, the average number of responses before reinforcement will be set by the schedule of 10.

Gambling machines, roulette wheels, and lotteries all work using variable ratios schedules. On a roulette wheel with 36 numbers, we know that we will win over time about once every 36 spins of the wheel. However, the outcome is unpredictable. We might just as often win immediately after our first bet or have to wait for many spins. It is this unpredictability that makes gambling exciting but also potentially addictive.

In interval schedules, the reinforcer is given following a response but only after an amount of time has elapsed since the last reinforcer, regardless of how many responses are made. The interval may be fixed or variable. If a coffee machine takes exactly a minute to brew the next cup before it can be dispensed, we will only get our drink by pushing the button once a minute. We don't get anything apart from frustration by pushing the dispense button repeatedly before that time. This is an example of an FI schedule of reinforcement. Many other types of more complex schedules have been developed, but we will not consider them here.

Slide 6

Why are reinforcement schedules important? One reason is that the choice of schedule has a marked effect on how long it takes to learn to make a specific response when presented with a specific stimulus. Second, it influences the intensity and pattern of responding over time. Finally, it affects the persistence of a learned behaviour, even after the reinforcer is withdrawn.

The standard way of measuring behaviour from operant boxes is a graph that plots the cumulative number of responses on the vertical axis against time on the horizontal axis. A tick mark on the curve or line shows where a reinforcer is given. The slope of the line indicates the rate of responding, with the steeper the line, the faster the rate. If the animal stops responding, the line becomes horizontal.

This figure shows performance on a variable ratio experiment for an animal that has already learned to respond to a stimulus to obtain a reward. The animal is responding at a fairly continuous rate of just under once a second, to get a reinforcer around once a minute. Early in learning, it is usually necessary to have a continuous reinforcement schedule or one with a high probability of

reinforcement. This makes sure that the initial trial and error response is reinforced and is likely to be repeated. Once learned, the ratio can be slowly increased across trials so that reinforcement happens less and less often while maintaining responding. This process is an example of shaping.

Slide 7
What does behaviour look like under different ratios of response to reinforcement? This first diagram shows the cumulative rate of responding over time under four different fixed ratio schedules-- reinforcement every 70, 120, 180, and 320 responses. The first thing to notice is the characteristic stepped or staircase shape of the record, with bursts of responding showing as the steep part of the trace until the reinforcement is obtained, at which point the response stops, and the trace shows a horizontal line. This same pattern is sitting whether we are measuring the behaviour of a pigeon, a rat, or human for money. It is a basic characteristic of fixed ratio schedules.

The duration of the pause and behaviour increases systematically with the ratio. The more response is needed to obtain a reward, the longer the pause after reinforcement before the animal resumes responding. Once started, however, the rate of responding is the same, whatever the ratio.

We see a similar pattern for fixed interval schedules, although the characteristic curve is rather more scallop shaped than a sharply defined staircase. Once the animal has learned the interval, they tend to stop or slow down the rate of responding after obtaining the predictably timed reward and start responding more quickly as the duration increases towards the end of the expected interval.

Slide 8
What does behaviour look like with variable schedules? This diagram shows the cumulative rate of responding over time and at different variable interval schedules-- reinforcement, on average, every 1, 2, 3, 6, and 10 seconds. The first thing to notice is that unlike the staircase of the fixed schedules, we see almost continuous behaviour, even after the animal has just obtained a reward. This is because there is uncertainty, although it is unlikely another reward may come immediately, so it is adaptive for the animal to keep responding just in case. However, we also see that the rate of responding is slower the longer the meeting interval between reinforcements or the greater the number of responses that have to be made, on average, to get the reward.

This makes good adaptive sense. If the reward is food in the wild, then the animal will adjust the amount of energy it expends by varying the rate of responding. This helps to make sure that there is a net gain in energy in over energy out. We don't have to assume that the animal is doing this consciously.

Slide 9
Next, let's look at what happens when a response stops being reinforced-- if the pigeon no longer gets a reward when it pecks a key. What happens is that eventually the pigeon stops pecking and ignores the stimulus. This is known as extinction. The pattern of extinction is different in fixed and variable schedules.

For fixed schedules, the reinforcement is predictable, so it is more obvious when an expected reward is not delivered. However, rather than stopping immediately, we typically see a continuation in behaviour. Indeed, the rate of responding may even increase before eventually stopping. This is sometimes called pre-extinction burst. Some behaviour will continue, but the bouts of behaviour become shorter and the intervals between them longer before stopping altogether.

With variable schedules we see behaviour continuing for much longer. This is because there is greater uncertainty about when the next reward is expected. The animal is less able to determine whether it is just an unusually large pause or whether the behaviour is no longer reinforced. The

longer the mean time or mean number of responsibles in the variable schedule, the longer the behaviour will continue before it starts to extinguish.

Variable ratio schedules show a similar stepped pattern of extinction to fixed ratio schedules, while variable interval schedules show a more continuous slowing of responding before it stops. This figure shows some of the characteristic extinction curves of the different types of schedules.

In general, intermittent or partial reinforcement, whether on a fixed or variable schedule, is more resistant to extinction than when an animal has been previously continuously reinforced every response. This is called the partial reinforcement extinction effect.

What does it mean that the animal stops responding after the behaviour is no longer reinforced? Does it mean that the animal has forgotten the association between the stimulus and response and subsequent reinforcer? If this was the case, then we would predict that the animal would take as long to relearn the association if the reinforcement contingency was introduced. This was not what we see. Instead, the previous behaviour is quickly reestablished and much more quickly than it took to learn originally.

Note, this is very similar to Ebbinghaus's use of saving scores as an index of learning with his nonsense syllables. So the animal had not forgotten the association. Simply, it was no longer useful or adaptive to respond to a stimulus that did not ever produce a reward. The learned association remained within the animals learned repertoire and was available to be used when the circumstances changed.

Slide 10

Before we leave the subjects of reinforcement, conditioning, and extinction for now, fans of the TV show Big Bang Theory may remember an episode in which Sheldon uses operant principles to condition Penny's behaviour. If you feel like it, take a chance later to review the clip that is found easily on the web. If you do, and once you've watched it, you might want to try the quiz questions to test your understanding of what we have just covered in this topic.