Chapter 5

Behavioural Process

The "Discovery" of Operant Conditioning
- Thorndike made use of puzzle boxes for cats, he then recorded how long it took the cat to escape from the box
- if there is a reward for escaping the box, the response would occur more often, vice verse. This is the law of effect. S-R association.
- Tolman: Stimulus -> Response -> Outcome
- Classical vs. Operant Conditioning
  - Classical conditioning: organisms experience an outcome (US) whether or not they have learned the CR.
  - Operant conditioning: the outcome O is wholly dependent on whether the organism performs the response R.
  - If outcome occurs regardless of responding, it is classical; if it is contingent on a response, then it's operant
  - Both show extinction
- Free-operant Learning
  - Discrete trail paradigms: the experimenter defined the beginning and the end of each trial
  - Free-operant: Skinner automated data collection, no more experimenter intervention. The animal now controls the rate of responding
  - Skinner cage, trough in one wall through which food can be delivered
  - Cumulative recorder: the height of the line at any given time represents the number of responses that have been made in the entire experiment up to that time

Components of the Learned Association
- Stimuli
  - Discriminative Stimuli: stimuli that signal whether a particular response will lead to a particular outcome
  - S (lights on) -> R (press lever) -> O (get food)
  - habit slip: rat ran through the maze and passed a pile of food without stopping
  - Protestant ethic effect: rats will often continue to work to obtain food in preference to eating identical food from a freely available source
  - above two effects show how strongly discriminative stimuli can evoke the associated responses, overruling other possible behaviours
- Response
  - organisms learn to make specific response in order to acquire their desired outcome
  - shaping: successive approximations to the desired response are reinforced
  - chaining: organisms are gradually trained to execute complicated sequences of discrete response
  - backwards chaining
- Reinforcers
  - a consequence of behaviour that leads to increased likelihood of that behaviour in the future
  - Primary reinforcer: organisms have innate drives to obtain these things and therefore to repeat behaviours that provide access to these things
  - Secondary reinforcers: no intrinsic value, but that have been paired with primary reinforcers
  - negative contrast: organisms given a less preferred reinforcer in place of an expected and preferred reinforcer will respond less strongly for the less reinforcer than if they had been given that less preferred reinforcer all along
- Punishers
  - punishments are not as effective as reinforcers in controlling behaviour
    - 1. discriminative stimuli for punishment can encourage cheating
    - 2. concurrent reinforcement can undermine the punishment
    - 3. punishment leads to more variable behaviour
    - 4. initial intensity matters

Putting it All together: Building the S-R-O Association
- Timing Affects Learning
  - most effective when the R-O interval is short
  - pre-commitment: to make a choice that is difficult to change later
- Outcomes can be added or subtracted
  - positive punishment, negative punishment, positive reinforcer, negative reinforcer
- Reinforcement need not follow every response
  - fixed-ratio: some fixed number of responses must be made before a reinforcer is delivered
    - e.g., must press lever 5 times to get food delivered
    - short break in responding after a reinforcement is called post-reinforcement
  - fixed-interval: reinforces the first response after a fixed amount of time
    - e.g., baking a cake without a timer
  - variable-ratio: reinforces after a certain average number of responses
    - e.g., gambling, slot machine
  - variable-interval: reinforces the first response after a particular time interval
    - e.g., checking facebook updates for a particular person
  - VR > FR > VI > FI

Choice Behaviour
- concurrent reinforcement schedule: the organism can make any of several possible responses, each leading to a different outcome

- variable-interval schedules and the matching law
  - matching law: given two responses that are reinforced on VI schedules, an organism's relative rate of making each response will match the relative rate of reinforcement for that response
  - $$\frac{\text{rate of response A}}{\text{rate of response B}} = \frac{\text{rate of VI reinforcement for A}}{\text{rate of VI reinforcement for B}}$$

- The Premack Principle: Responses as Reinforcers
  - in both rats and children, the opportunity to perform a highly frequent behaviour can reinforce a less frequent behaviour
  - response deprivation hypothesis: the critical variable is not which response is normally more frequent but merely which response has been restricted: by restricting the ability to execute almost any response, you can make the opportunity to perform that response reinforcing.

Brain Substrates

The Dorsal Striatum and Stimulus-Response Learning
- basal ganglia: Information from the sensory cortex to the motor cortex can also travel via a indirect route, through the basal ganglia. Which is a collection of ganglia that lie at the base of the forebrain. It contains a part called the dorsal striatum, which receives highly processed stimulus information from sensory cortical areas and projects to the motor cortex, which produces a behavioural response.
  - rats with lesion of the dorsal striatum can learn a simple R-O association, as well as control rats can. But cannot learn if a discriminative stimuli are involved

The Orbitofrontal Cortex and Learning to Predict Outcomes
- prefrontal cortex and orbitofrontal cortex: appears to contribute to goal-directed behaviour by representing predicted outcomes
- orbitofrontal cortex receives inputs conveying the full range of sensory modalities and also visceral sensations, allowing this brain area to integrate many types of information
- volunteers ate chocolate, ones who report that they really liked the chocolate had heightened activity in the medial area of the orbitofrontal cortex, but those who reported satiety had heightened activity in the lateral orbitofrontal cortex.
- some neurons also code to the expected outcome

Mechanisms of Reinforcement Signaling in the Brain
- "wanting" and "liking" the brain
  - ventral tegmental area (VTA): a small region in the midbrain of rats, humans, and other mammals, and its stimulation is a powerful reinforcer, also known as "pleasure centers"
  - researchers believe that we have separated brain systems for signaling hedonic value, meaning the goodness of a reinforcer, or how much we like it, that are distinct from those signaling motivational value, meaning how much we want a reinforcer and how hard we are wiling to work to obtain it
  - only when both "like" and "want" are present will the arrival of the reinforcer evoke responding and strengthen the S-R assocation
- Dopamine: How the brain signals "wanting"?
- Endogenous Opioids: how the brain signals "liking"?
  - endogenous opioids: naturally occurring neurotransmitter-like substance with many of the same effects as opiate drugs. When released to the body, they lessens the normal perception of pain and producing feelings of euphoria
  - activate same portion of the brain as heroin and morphine

lesions of the dorsal striatum interferes with S-R connections, but not R-O connection
lesions of the orbitofrontal cortex shows inflexible or inappropriate responding, the cortex usually plays a role in predicting the outcome of responses

Clinical Perspective

- pathological addiction: a strong habit, that is maintained despite known harmful consequences
- drug addiction results from several factors, including positive reinforcement (the pleasurable high), negative reinforcement (avoiding withdrawal), and malfunction of the dopaminergic "wanting" system (the craving)
- amphetamine causes dopaminergic neurons to release higher levels of dopamine, cocaine blocks dopamine reuptake, so that dopamine remains in the synapse longer before being reabsorbed. Both are to increase amount of dopamine in the system
- most treatment plans for addiction include cognitive therapy, often centered on self-help sessions with a support group.
  - addictions are considered a strong S-R-O association, the challenge of the treatment is to break or reduce the strength of the conditioned association, or extinction. Refer to the effectiveness of naltrexone
  - Distancing: avoiding the stimuli that trigger the unwanted response.
  - Reinforcement of Alternate behaviours: if the smoker makes it through a whole week without a cigarette, she can reinforce her own behaviour by treating herself to a favourite food or activity
  - Delayed reinforcement: whenever the smoker gets the urge to light up, she can impose a fixed delay before giving in to it.
- Most effective when combining cognitive, behaviour, and medical treatments

Chapter 6 - Generalization and Discrimination Learning

- Discrimination: the perception of differences between stimuli

- Generalization gradient: a curve showing how changes in the physical properties of the stimuli correspond to changes in responding. After training for a single stimulus has been reinforced repeatedly, generalization gradients around that trained stimulus show a peak.
- consequential region: identify the set of all stimuli that have the same consequences as the training stimulus. suggests that amimals consistently expect that the chance that two stimuli will have the same consequence drops off sharply as the stimuli become more distinct
- Discrete-component representation: each possible stimulus is represented by its own unique node in the model, applicable to situations in which the similarity between cues is small enough that there is negligible transfer from one to another. It fails in cases where stimuli have a high degree of physical similarity, since the model then produce unrealistic generalization gradients.
- Discrimination training: people learning to increase their sensitivity to subtle variations in stimuli