Operant conditioning

4.1 Introduction to Operant conditioning (or **instrumental conditioning**) is a type of learning in which an individual's behavior is modified by its antecedents and consequences. Instrumental conditioning was first discovered and published by Jerzy Konorski and was also referred to as *Type II reflexes*. Mechanisms of instrumental conditioning suggest that the behavior may change in form, frequency, or strength. The expressions "operant behavior" and "respondent behavior" were popularized by B. F. Skinner who worked on reproduction of Konorski's experiments. The former refers to "an item of behavior that is initially spontaneous, rather than a response to a prior stimulus, but whose consequences may reinforce or inhibit recurrence of that behavior".

Operant conditioning is distinguished from *classical conditioning* (or *respondent conditioning*) in that operant conditioning deals with the reinforcement and punishment to change behavior. Operant behavior operates on the environment and is maintained by its antecedents and consequences, while classical conditioning is maintained by conditioning of reflexive (reflex) behaviors, which are elicited by antecedent conditions. Behaviors conditioned through a classical conditioning procedure are not maintained by consequences. They both, however, form the core of behavior analysis and have grown into professional practices.

Historical notes

Thorndike's law of effect

Operant conditioning, sometimes called *instrumental learning*, was first extensively studied by Jerzy Konorski and next by Edward L. Thorndike (1874–1949), who observed the behavior of cats trying to escape from home-made puzzle boxes.[4] When first constrained in the boxes, the cats took a long time to escape. With experience, ineffective responses occurred less frequently and successful responses occurred more frequently, enabling the cats to escape in less time over successive trials. In his law of effect, Thorndike theorized that behaviors followed by satisfying consequences tend to be repeated and those that produce unpleasant consequences are less likely to be repeated. In short, some consequences *strengthened* behavior and some consequences *weakened* behavior. Thorndike produced the first known animal learning curves through this procedure.

Skinner

B.F. Skinner (1904–1990) is the person whose work is most often cited in connection with operant conditioning. His book "The Behavior of Organisms", published in 1938, initiated his lifelong study of operant conditioning and its application to human and animal behavior. Following the ideas of Ernst Mach, Skinner rejected Thorndike's reference to unobservable mental states such as satisfaction, building his analysis on observable behavior and its equally observable consequences.

To implement his empirical approach, Skinner invented the operant conditioning chamber in which subjects such as pigeons and rats were isolated from extraneous stimuli and free to make one or two simple, repeatable responses. Another invention, the cumulative recorder, produced a graphical record of these responses from which response rates could be estimated. These records were the primary data that Skinner and his colleagues used to explore the effects on response rate of various reinforcement schedules. A reinforcement schedule may be defined as "any procedure that delivers a reinforcer to an organism according to some well-defined rule". The effects of schedules became, in turn, the basic experimental data from which Skinner developed his account of operant conditioning. He also drew on many less formal observations of human and animal behavior.

Many of Skinner's writings are devoted to the application of operant conditioning to human behavior. In 1957, Skinner published *Verbal Behavior*, which extended the principles of operant conditioning to language, a form of human behavior that had previously been analyzed quite differently by linguists and others. Skinner defined new functional relationships such as "mands" and "tacts" to capture the essentials of language, but he introduced no new principles, treating verbal behavior like any other behavior controlled by its consequences, which included the reactions of the speaker's audience.

Consequences that shape behavior: Reinforcement, punishment, and extinction

Reinforcement and punishment are the core tools of operant conditioning. It is important to realise that some terminology in operant conditioning is used in a way that is different from everyday use.

1. **Positive** means a stimulus is delivered following a response
2. **Negative** means a stimulus is withdrawn following a response

3. **Reinforcement** is a consequence that causes a behavior to occur with greater frequency.
4. **Punishment** is a consequence that causes a behavior to occur with less frequency.

There are an additional fifth and sixth procedures

1. **Antecedent stimuli** (Precede): Occurs before a behavior happens.
2. **Extinction** is caused by the lack of any consequence following a behavior. When a behavior is inconsequential (i.e., producing neither favorable nor unfavorable consequences) it will occur less frequently. When a previously reinforced behavior is no longer reinforced with either positive or negative reinforcement, it leads to a decline (extinction) in that behavior.

This creates a total of five basic consequences -

1. **Positive reinforcement** (Reinforcement): Occurs when a behavior (response) is followed by a stimulus that is appetitive or rewarding, increasing the frequency of that behavior. In the Skinner box experiment, a stimulus such as food or a sugar solution can be delivered when the rat engages in a target behavior, such as pressing a lever. This procedure is usually called simply *reinforcement*.
2. **Negative reinforcement** (Escape): Occurs when a behavior (response) is followed by the removal of an aversive stimulus, thereby increasing that behavior's frequency. In the Skinner box experiment, negative reinforcement can be a loud noise continuously sounding inside the rat's cage until it engages in the target behavior, such as pressing a lever, upon which the loud noise is removed.
3. **Positive punishment** (Punishment) (also called "Punishment by contingent stimulation"): Occurs when a behavior (response) is followed by a stimulus, such as introducing a shock or loud noise, resulting in a decrease in that behavior. *Positive punishment* is sometimes a confusing term, as it denotes the "addition" of a stimulus or increase in the intensity of a stimulus that is aversive (such as spanking or an electric shock). This procedure is usually called simply *punishment*.
4. **Negative punishment** (Penalty) (also called "Punishment by contingent withdrawal"): Occurs when a behavior (response) is followed by the removal of a stimulus, such as taking away a child's toy following an undesired behavior, resulting in a decrease in that behavior.

5. **Extinction**: Occurs when a behavior (response) that had previously been reinforced is no longer effective. For example, a rat is first given food many times for lever presses. Then, in "extinction", no food is given. Typically the rat continues to press more and more slowly and eventually stops, at which time lever pressing is said to be "extinguished."

Some other common terms and procedures

- **Escape and Avoidance** In escape learning, a behavior terminates an (aversive) stimulus. For example, shielding one's eyes from sunlight terminates the (aversive) stimulation of bright light in one's eyes. In avoidance learning, the behavior precedes and prevents an (aversive) stimulus, for example putting on sun glasses before going outdoors. Because, in avoidance, the stimulation does not occur, avoidance behavior seems to have no means of reinforcement. Indeed this non-occurrence of the stimulus has been a problem for reinforcement theory, which has been dealt with in various ways. See section on avoidance learning below.
- **Noncontingent reinforcement** refers to delivery of reinforcing stimuli regardless of the organism's behavior. Noncontingent reinforcement may be used in an attempt to reduce an undesired target behavior by reinforcing multiple alternative responses while extinguishing the target response. As no measured behavior is identified as being strengthened, there is controversy surrounding the use of the term noncontingent "reinforcement".

- **Schedules of reinforcement** Schedules of reinforcement are rules that control the delivery of reinforcement. The rules specify either the time that reinforcement is to be made available, or the number of responses to be made, or both.
    - Fixed interval schedule: Reinforcement occurs following the first response after a fixed time has elapsed after the previous reinforcement.
    - Variable interval schedule: Reinforcement occurs following the first response after a variable time has elapsed from the previous reinforcement.
    - Fixed ratio schedule: Reinforcement occurs after a fixed number of responses have been emitted since the previous reinforcement.
    - Variable ratio schedule: Reinforcement occurs after a variable number of responses have been emitted since the previous reinforcement.
    - Continuous reinforcement: Reinforcement occurs after each response.

- **Discrimination, generalization and context**. Most behavior is under stimulus control. Several aspects of this may be distinguished:
    - "Discrimination" typically occurs when a response is reinforced only in the presence of a specific stimulus. For example, a pigeon might be fed for pecking at a red light and not at a green light; in consequence, it pecks at red and stops pecking at green. Many complex combinations of stimuli and other conditions have been studied; for example an organism might be reinforced on an interval schedule in the presence of one stimulus and on a ratio schedule in the presence of another.
    - "Generalization" is the tendency to respond to stimuli that are similar to a previously trained discriminative stimulus. For example, having been trained to peck at "red" a pigeon might also peck at "pink", though usually less strongly.
    - "Context" refers to stimuli that are continuously present in a situation, like the walls, tables, chairs, etc. in a room, or the interior of an operant conditioning chamber. Context stimuli may come to control behavior as do discriminative stimuli, though usually more weakly. Behaviors learned in one context may be absent, or altered, in another. This may cause difficulties for behavioral therapy, because behaviors learned in the therapeutic setting may fail to occur elsewhere.

4.2 Operant conditioning to change human behavior

Researchers have found the following protocol to be effective when they use the tools of operant conditioning to modify human behavior:

1. **State goal** (aims for the study) That is, clarify exactly what changes are to be brought about. For example, "reduce weight by 30 pounds."
2. **Monitor behavior** (log conditions) Keep track of behavior so that one can see whether the desired effects are occurring. For example, keep a chart of daily weights.
3. **Reinforce desired behavior** (give reward for proper behavior) For example, congratulate the individual on weight losses. With humans, a record of behavior may serve as a reinforcement. For example, when a participant sees a pattern of weight loss, this may reinforce continuance in a behavioral weight-loss program. A more general plan is the token economy, an exchange system in which tokens are given as rewards for desired behaviors. Tokens may later be exchanged for a desired prize or rewards such as power, prestige, goods or services.

4. **Reduce incentives to perform undesirable behavior** For example, remove candy and fatty snacks from kitchen shelves.

Factors that alter the effectiveness of consequences

When using consequences to modify a response, the effectiveness of a consequence can be increased or decreased by various factors. These factors can apply to either reinforcing or punishing consequences.

1. **Satiation/Deprivation:** The effectiveness of a consequence will be reduced if the individual's "appetite" for that source of stimulation has been satisfied. Inversely, the effectiveness of a consequence will increase as the individual becomes deprived of that stimulus. If someone is not hungry, food will not be an effective reinforcer for behavior. Satiation is generally only a potential problem with primary reinforcers, those that do not need to be learned such as food and water.
2. **Immediacy:** After a response, how immediately a consequence is then felt determines the effectiveness of the consequence. More immediate feedback will be more effective than less immediate feedback. If someone's license plate is caught by a traffic camera for speeding and they receive a speeding ticket in the mail a week later, this consequence will not be very effective against speeding. But if someone is speeding and is caught in the act by an officer who pulls them over, then their speeding behavior is more likely to be affected.
3. **Contingency:** If a consequence does not contingently (reliably, or consistently) follow the target response, its effectiveness upon the response is reduced. But if a consequence follows the response consistently after successive instances, its ability to modify the response is increased. The schedule of reinforcement, when consistent, leads to faster learning. When the schedule is variable the learning is slower. Extinction is more difficult when learning occurs during intermittent reinforcement and more easily extinguished when learning occurs during a highly consistent schedule.
4. **Size:** This is a "cost-benefit" determinant of whether a consequence will be effective. If the size, or amount, of the consequence is large enough to be worth the effort, the consequence will be more effective upon the behavior. An unusually large lottery jackpot, for example, might be enough to get someone to buy a one-dollar lottery ticket (or even buying multiple tickets). But if a lottery jackpot is small, the same person might not feel it to be worth the effort of driving out and finding a place to buy a ticket. In this example, it's also useful to note that "effort" is a punishing consequence. How these

opposing expected consequences (reinforcing and punishing) balance out will determine whether the behavior is performed or not.

Operant variability

Operant variability is what allows a response to adapt to new situations. Operant behavior is distinguished from reflexes in that its **response topography** (the form of the response) is subject to slight variations from one performance to another. These slight variations can include small differences in the specific motions involved, differences in the amount of force applied, and small changes in the timing of the response. If a subject's history of reinforcement is consistent, such variations will remain stable because the same successful variations are more likely to be reinforced than less successful variations. However, behavioral variability can also be altered when subjected to certain controlling variables.

Avoidance learning

In avoidance learning an organism's behavior is reinforced by the termination or prevention of an (assumed aversive) stimulus. There are two kinds of commonly used experimental settings: discriminated and free-operant avoidance learning.

Discriminated avoidance learning

In discriminated avoidance learning, a novel stimulus such as a light or a tone is followed by an aversive stimulus such as a shock (CS-US, similar to classical conditioning). During the first trials (called escape-trials) the animal usually experiences both the CS (Conditioned Stimulus) and the US (Unconditioned Stimulus), showing the operant response to terminate the aversive US. During later trials, the animal will learn to perform the response during the presentation of the CS thus preventing the aversive US from occurring. Such trials are called "avoidance trials."

Free-operant avoidance learning

In this experimental session, no discrete stimulus is used to signal the occurrence of the aversive stimulus. Rather, the aversive stimulus (mostly shocks) are presented without explicit warning stimuli. There are two crucial time intervals determining the rate of avoidance learning. This first one is called the S-S-interval (shock-shock-interval). This is the amount of time which passes during successive presentations of the shock (unless the operant response is performed). The other one is called the R-S-interval (response-shock-interval) which specifies the length

of the time interval following an operant response during which no shocks will be delivered. Note that each time the organism performs the operant response, the R-S-interval without shocks begins anew.

Two-process theory of avoidance

This theory was originally proposed in order to explain discriminated avoidance learning, in which an organism learns to avoid an aversive stimulus by escaping from a signal for that stimulus. The theory assumes that two processes take place:

a) Classical conditioning of fear.

During the first trials of the training, the organism experiences the pairing of a CS with an aversive US. The theory assumes that during these trials an association develops between the CS and the US through classical conditioning and, because of the aversive nature of the US, the CS comes to elicit a conditioned emotional reaction (CER) – "fear."

b) Reinforcement of the operant response by fear-reduction.

As a result of the first process, the CS now signals fear; this unpleasant emotional reaction serves to motivate operant responses, and those responses that terminate the CS are reinforced by fear termination. Although, after this training, the organism no longer experiences the aversive US, the term "avoidance" may be something of a misnomer, because the theory does not say that the organism "avoids" the US in the sense of anticipating it, but rather that the organism "escapes" an aversive internal state that is caused by the CS.

Operant conditioning in economics

Both psychologists and economists have become interested in applications of operant conditioning concepts and findings to the behavior of humans in the marketplace. One concept that encompasses both of economics and instrumental conditioning is consumer demand. With consumer demand, the focus is on the price of the commodity and the amount purchased. The degree to which price influences consumption is defined as being the elasticity of demand. Certain commodities are more elastic than others. Price change in certain foods can affect the amount bought, while gasoline and essentials seem to be less effected by price changes. For these examples, gasoline and essentials would be less elastic than certain foods like cake and candy. On a graph model representation, something less elastic would not be stretched out as far as a commodity that's consumption fluctuates greatly due to the price.

Questions about the law of effect

A number of observations seem to show that operant behavior can be established without reinforcement in the sense defined above. Most cited is the phenomenon of autoshaping (sometimes called "sign tracking"), in which a stimulus is repeatedly followed by reinforcement, and in consequence the animal begins to respond to the stimulus. For example, a response key is lighted and then food is presented. When this is repeated a few times a pigeon subject begins to peck the key even though food comes whether the bird pecks or not. Similarly, rats begin to handle small objects, such as a lever, when food is presented nearby. Strikingly, pigeons and rats persist in this behavior even when pecking the key or pressing the lever leads to less food (omission training).

These observations and others appear to contradict the law of effect, and they have prompted some researchers to propose new conceptualizations of operant reinforcement (e.g. A more general view is that autoshaping is an instance of classical conditioning; the autoshaping procedure has, in fact, become one of the most common ways to measure classical conditioning. In this view, many behaviors can be influenced by both classical contingencies (stimulus-reinforcement) and operant contingencies (response-reinforcement), and the experimenter's task is to work out how these interact.