

Operant Conditioning L-3

①

Sometimes called instrumental conditioning / Reward Learning

- Learning the r/ship b/w one's actions and their consequences
- Operant behaviours are controlled by consequences

It is the method of Learning that occur through rewards and punishment for behaviour, through Operant Conditioning

An association is made b/w the Behavior and the consequence for the behaviour.

→ Operant Behav coined by Skinner

Skinner used the term operant to refer to any active behaviour that operates upon the environment to generate consequences.

→ O.C is based on Law of effect given by Thorndike. which states that behavior that are rewarded are repeated and the (B^r) that are punished are avoided.

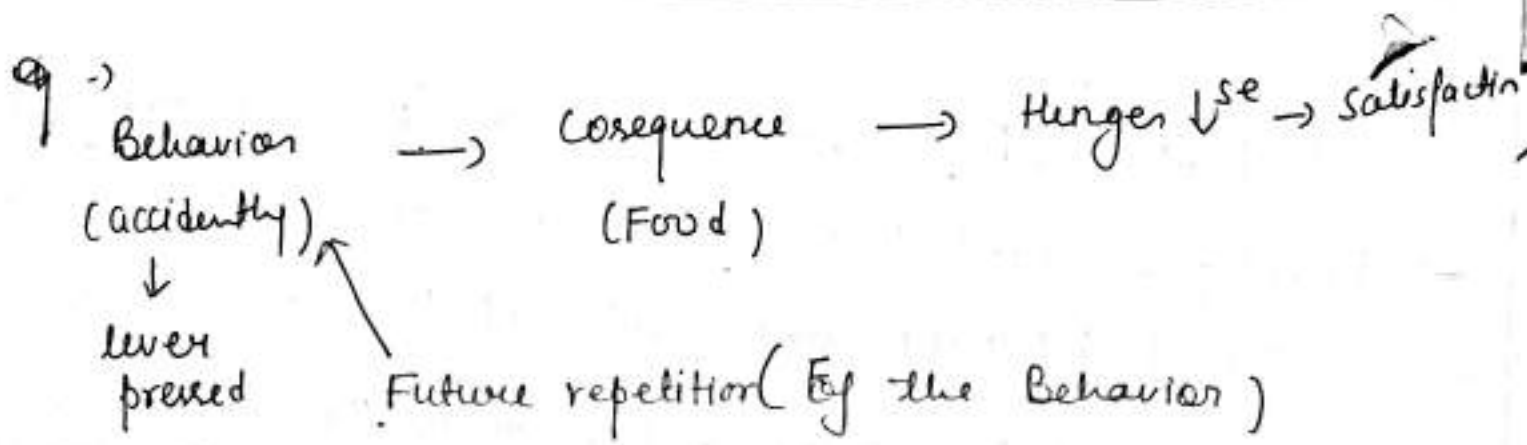
→ In OC is a procedure wherein learner perform some action which brings about certain changes in the envmt.

have →

i.e ↓ follow
Consequences

Influence likelihood of occurrence or non-occurrence of that behaviour or action.

∴ Learner operate on envmt → Produce consequence → ↑ or ↓ respons



Reinforcement

— Any event or condition that is the consequence of a response & that ↑^{se} the likelihood of the response being repeated.

+ve

↓
Any pleasant event or condition that follows the response and ↑^{se} the probability of response being repeated.

-ve. →

the probability that the desired response will be performed is ↑^{sed} by taking away or preventing something undesirable whenever the desired response is made.

Negative reinforcement = turned off once desired response has been achieved.

2 types

OC #2

Escape



Behaviour removes something undesirable

eg -> loud annoying buzzer tells you that, you that have not fastened ~~the~~ your car seat belt is not fastened can be taken away by fastening your seat belt.

Avoidance



The organism gets warning that an aversive stimulus will soon occur, and appropriate response completely avoids the aversive stimulus.

eg Paying Tax on time.



Value behaviour stops an aversive stimulus from ever happening.

weather +ve or -ve
Behaviour is being affected: It is more likely that you will, in the future, perform the same behaviour under similar circumstances.

*(Add on)

Punishment



giving something undesirable.

probability that a response will be made is ↓ sed by giving the organism something undesirable whenever the response is made.
eg Sending a child to his or her room (something undesirable) becoz he/she wrote on walls, = ↓ se in behaviour.

-ve
(Reinforcement)

→ taking away or preventing smthg undesirable.

When stimulus is ^(giving) applied, it is punishment, when removed ↓ taking away it is -ve reinforcement.

Reinforcement

* Add on

Primary



- any event or condition which naturally ↑ the likelihood of response
- No previous training required in order to produce effect

eg → $\frac{+ve}{(Food)}$ $\frac{-ve}{(Shock)}$

Secondary



- Do not work naturally
- Learner must have experience i.e. they get reinforcing property after consistent pairing with primary reinforcement
- Undergo extinction if presented alone w/o pairing
eg Cash prize / Medal

Factors that Influence Punishment

- 1) Strength = Stronger punishment = More effective
Pblm →
 - Ethically wrong
 - May result in conditioned fear
- 2) Should not be delivered in an arbitrary manner
- 3) " be consistent
- 4) Punishment should follow the reinforcement of desired behavior.
- 5) Indi must know why he got the punishment.
It help indi to evaluate it as a fair treatment
- 6) Contiguity → Punishment should be delivered immediately after undesirable (B)
as Time ↑ = Poor association b/w

Goal of Punishment) → ↓ se prob of (B^r)
" " -v reinf → ↑ se prob of B^r.

(Extinction, Generalisation, discrimination Same as CC,
↓
→ when light on
→ Animal peck = food
→ No light = No peck = No food.

Behavior that used to bring reward no longer does so.

Trained animal to peck for food when a green light is on.

After training → If it peck on illumination of any light,

Partial Reinforcement effect →

Rat A → every time lever press = Food
Rat B → Every other time = Food } both press lever.

↓
eg Food after 5 times lever press / w/o fixed criteria

Extinction training → RAT A show faster extinction.
RAT B → slowest extinction becoz it received occasional reinforcement.

How to deliver Reinforcement

Simple Schedule

Single type of reinforcement contingently maintained, within constant parameters are set is enforced, throughout the exptal session.

Compound.

Continuous Reinforcement

Every occurrence of the appropriate behaviour is reinforced
i.e. every correct response during acquisition is rewarded.

Partial Reinforcement

Set up in term of Time Elapsed from the delivery of last reinforcement.

Interval (Time)

Fixed Interval

After 10 minutes only irrespective of how fast or how many times you pressed the lever.

But organism gradually learns that it will get ~~see~~ reward after ~~see~~ the passage of specific reward.

which leads to condition known as Fixed Interval

Scallop → i.e. responses just before the time of reinforcement.

Variable Interval

Reinforcement is delivered at different variable time elapsing interval. i.e.

May get after 10 min next time after 1 min than 25 min.

(It produce High steady rate of responding.)

(Reinforcement is contingent upon the no. of responses Ratio emitted by organism as measured from last reinforcement.)

FR

VR

4 types of Partial Reinforcement.

1) Fixed Ratio → Behavior reinforced after a fixed number of responses
eg Piecemeal.

2) Variable Ratio (times you see how many) → Behavior will be reinforced after varying number of responses

3) Fixed Interval → Behavior will be reinforced for the first response after a fixed period of time has elapsed since the last reinforcement.
eg going to office to pick up bimonthly pay check

4) Variable Interval → Behavior will be reinforced for the first response after a varying period has elapsed since the last ~~reinforcement~~ reinforcement
eg Parent responding to crying child
(Time)



FR → In this Schedule of reinforcement Reinforcement is delivered after a Fixed no. of response had been performed.

no resp remain same on every trial
eg After 10 correct response.
eg

In both FR and FI → the rewarded response is followed by a Depression in the rate of Responding. Which is known as Post Reinforcement Pause.

Reason is not clear.
Only speculation ~~that~~ are those that animals may have learned that the responses immediately following the rewarded responses are rarely rewarded.

Scallop is missing in FR

VR → R^{mt} is delivered after various no. of at varied ratio

eg 1st time → 10 times
5 times
2 times. } varies from trial to trial

This provide maximum resistance to extinction.

VR > ~~FR~~ FR > VI > FI