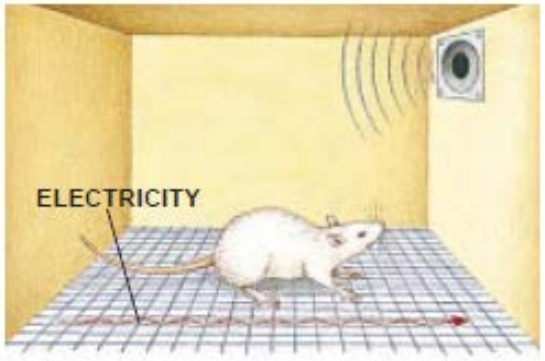
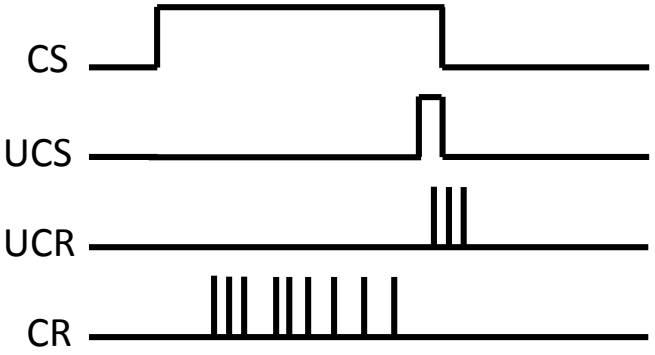
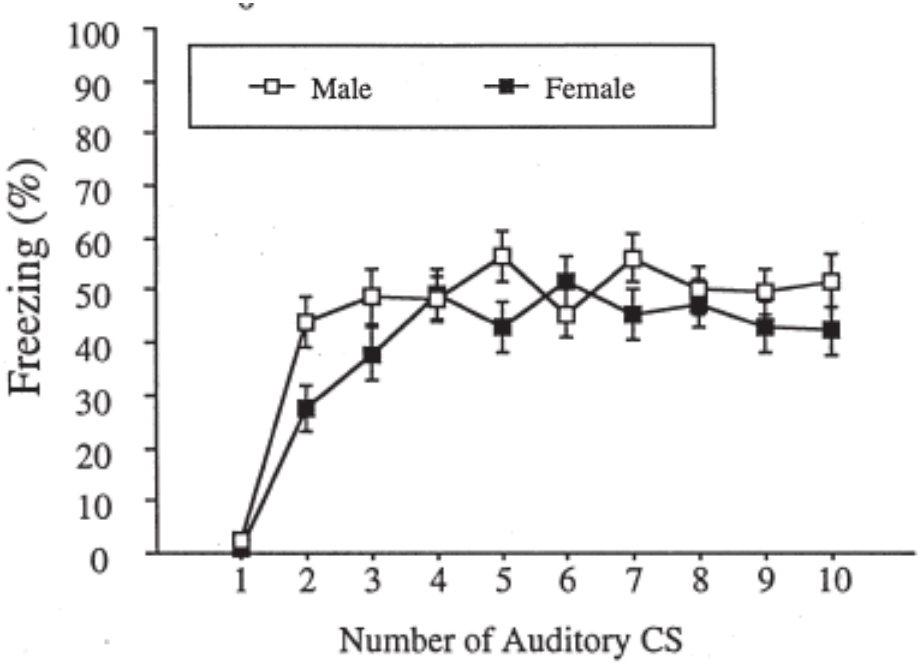
Learning and memory: Operant Conditioning and Goal-directed behaviour

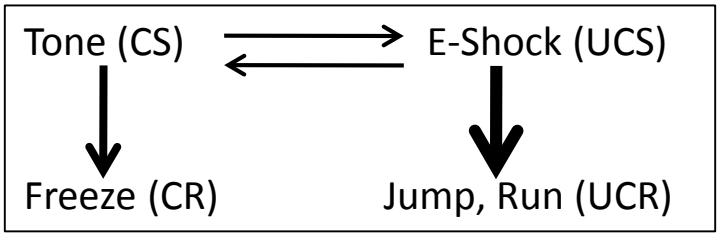
- Instrumental conditioning
- Operant Stimulus (S)
- Operant response (R)
- S-R association
- Law of Effect
- Unconditioned stimulus (US, UCS) = Reinforcer
- R-UCS association / R-Outcome association (R-O)
- Outcome-R association (O-R)
- Expectancy, Goal
- Reinforcement schedules
- Un/controllability
- Learned helplessness / Learned irrelevance
- Event (E1) → Event (E2) association

Classical conditioning or stimulus-stimulus conditioning: the learning curve for conditioned fear-freezing



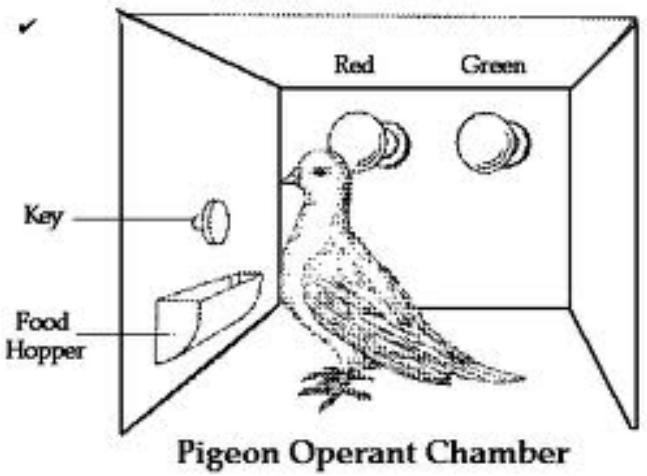




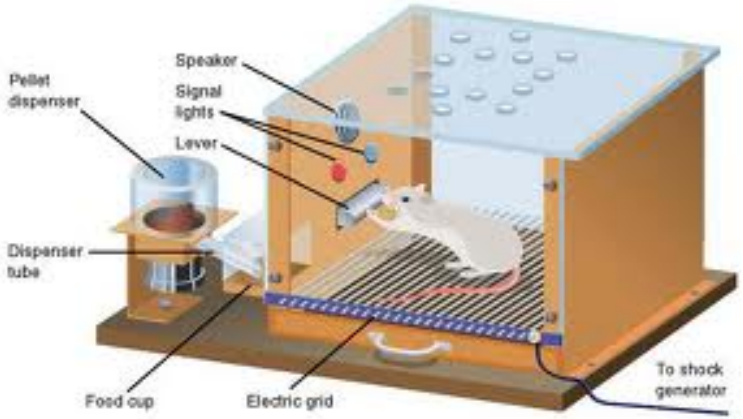


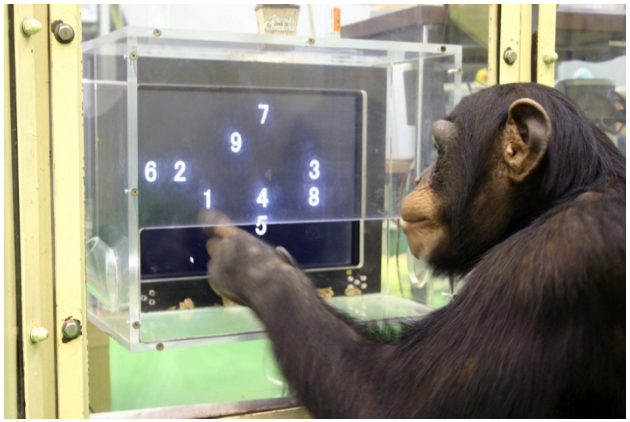
CS-UCS
Yes!

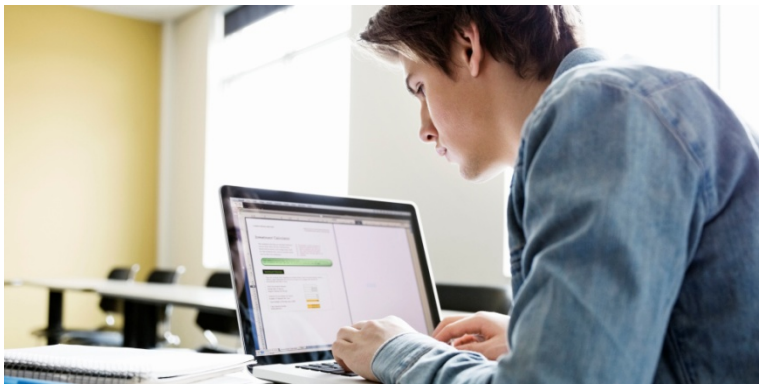
Operant Conditioning



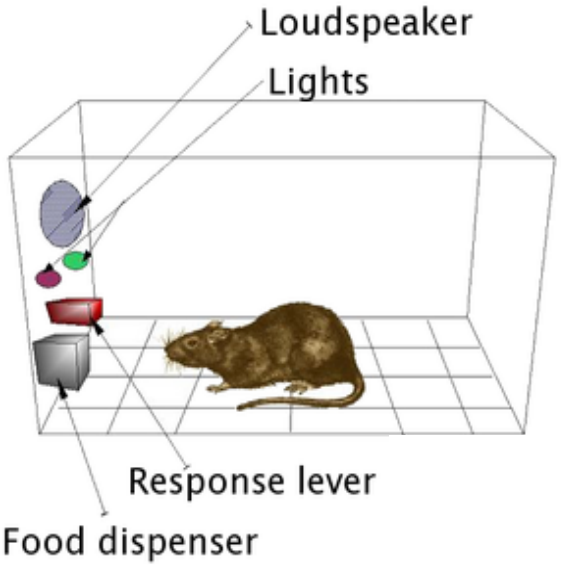
Pigeon Operant Chamber



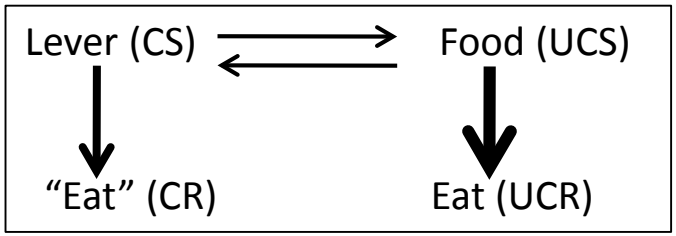




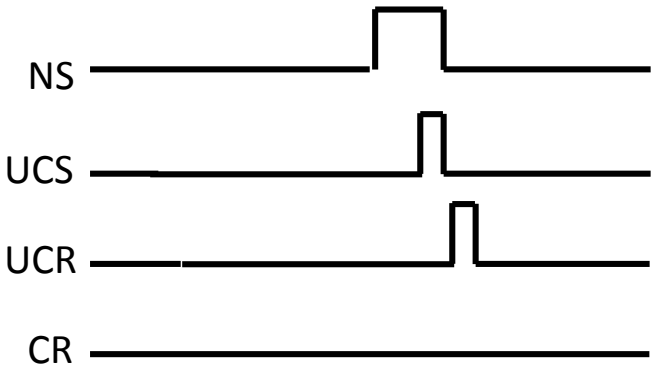
Can CS-US (classical) conditioning explain operant conditioning?



operant cond is not a special form of classical cond



Trial 1



only theoretically possible,
but does not happen in reality

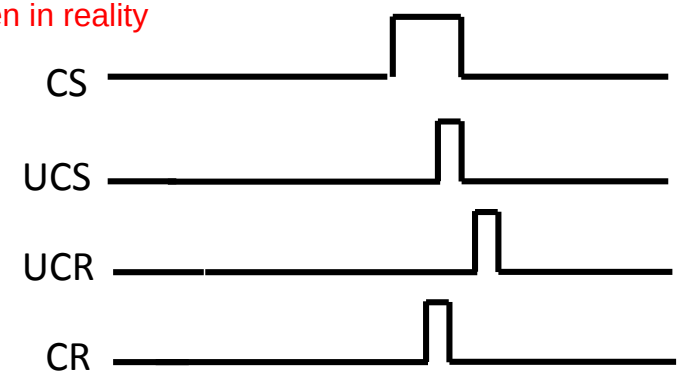
Lever

Food

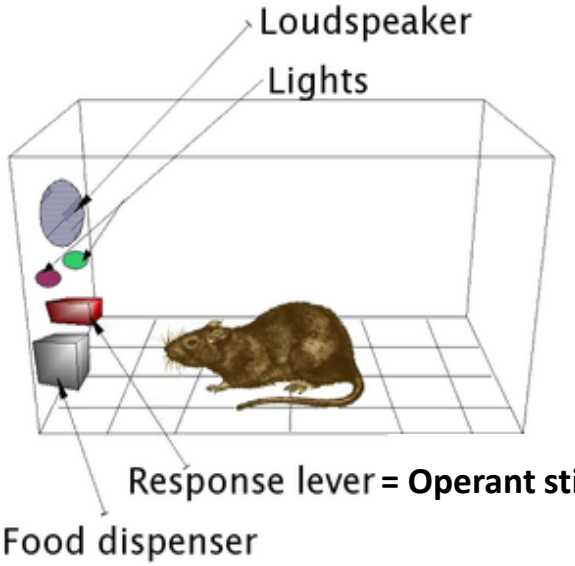
Approach, Eat

Approach, Lick, Touch, Bite Lever

Trial 10

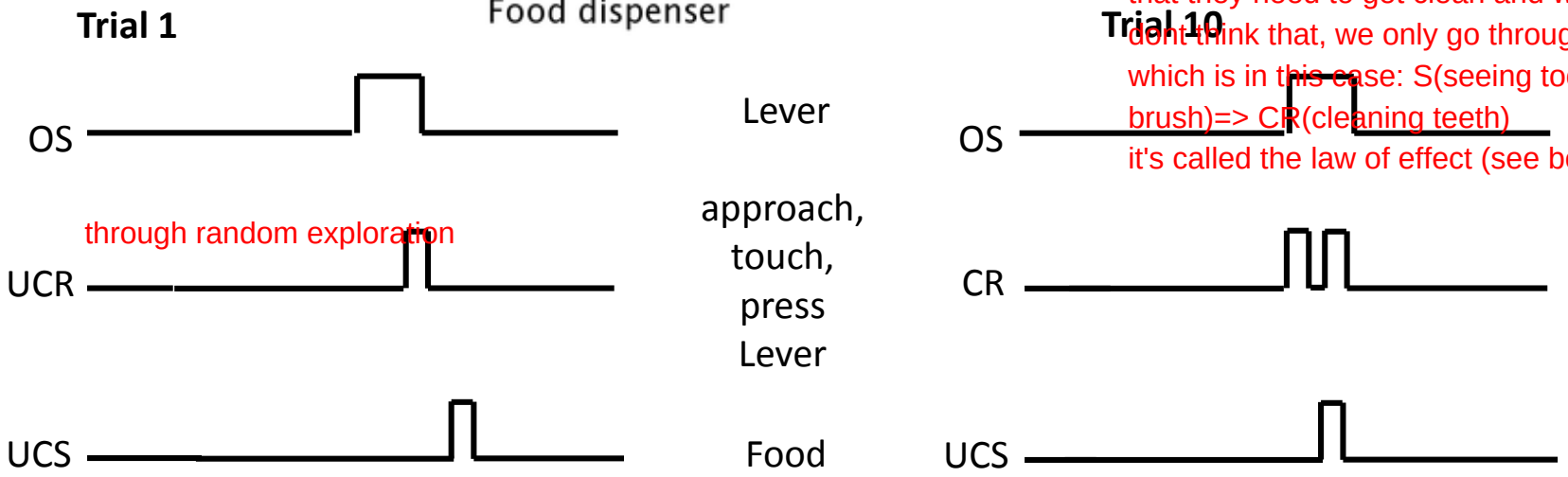


Can Stimulus-Response association explain operant conditioning?



when animals learn to demonstrate behaviour resulting into a reward, they do not perform beha because it results in reward, but they do it because the learnt S->CR
they don't show intelligent beha but they do S->CR

human ex: I brush teeth because then they are clean. In reality, we dont really do think that they need to get clean and white, we dont think that, we only go through a habit which is in this case: S(seeing tooth brush)=> CR(cleaning teeth)
it's called the law of effect (see below)



what ahppens after CR is not relevant for the animal, it has simply learnt the relationship S->CR

(O)perant (S)timulus → (UCR)esponse → Reinforcer (UCS) → (S)timulus → (CR)esponse

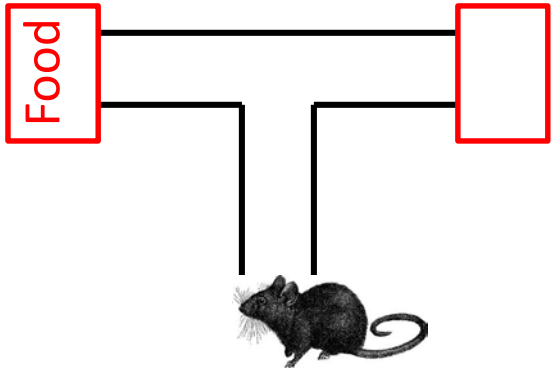
this is not the full explanation though

Law of Effect

Many experiments are not compatible with the Law of Effect (S-R)

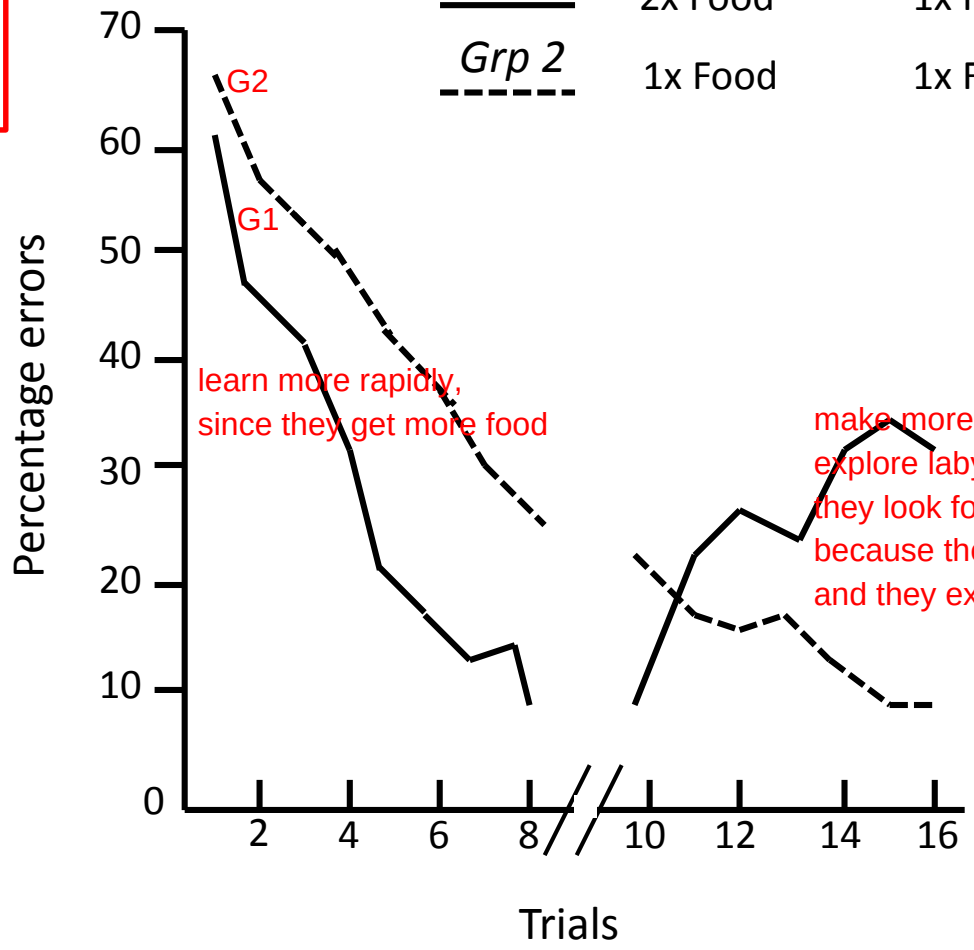
R: turning left or right

Rats in a multi-choice maze (labyrinth) with one route to a reward chamber



more UCS makes us learn faster

	Trials 1-9	Trials 10-16
Grp 1	2x Food	1x Food
Grp 2	1x Food	1x Food

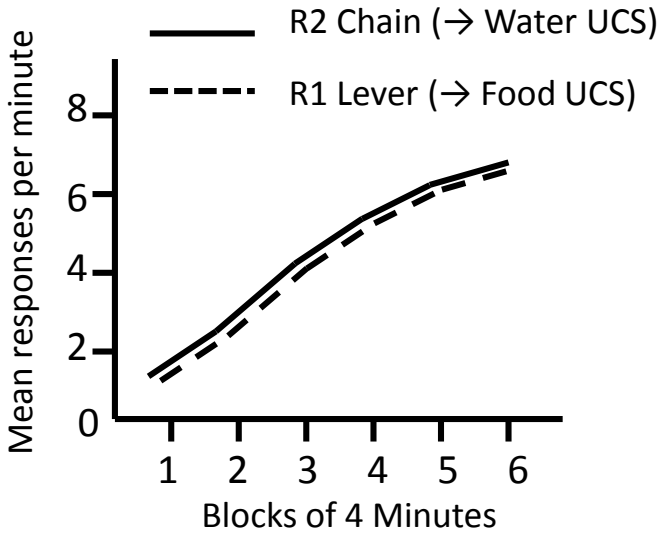


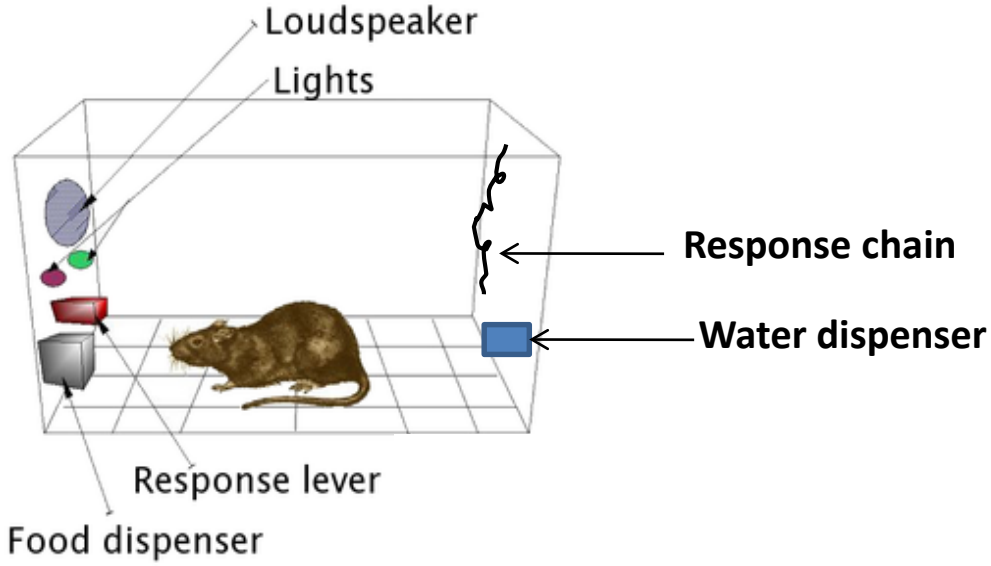
learn more rapidly, since they get more food

make more errors, because they explore labyrinth again, because they look for 2 foods not for 1 food, because their initial UCS is 2 foods and they expect 2 foods

Findings suggest that animals had some expectancy of the reward they would receive !

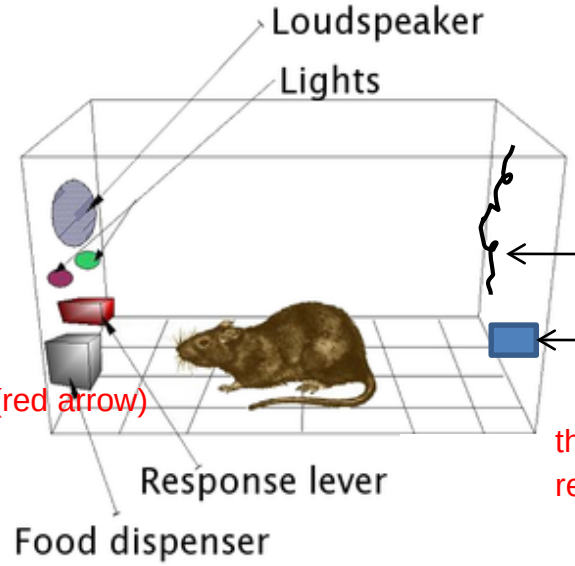
Can Response-UCS (expectancy) association explain operant conditioning?





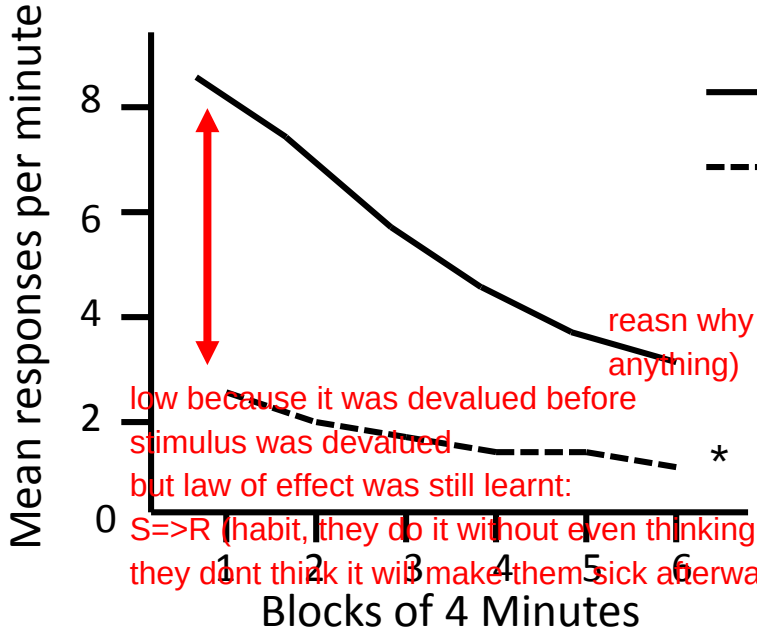
Training	*Devaluation*	*Test*
Lever R1→ Food UCS1	pair reward with a negative consequnece: aversive taste conditioning only works with food it does not know very well (something it does not get normally)	
Chain R2 → Water UCS2	**Food UCS1 → LiCl (Poison)** makes them feel sick (belly ache)	**R1 vs R2 (no UCS)** here they get neither food nor water but they do a response (see next slide)

Response-UCS (expectancy) association can explain operant conditioning



S->R would not explain a difference (red arrow)

the animal has a goal in mind expecting to be rewarded



reasn why it decreases is we test animals under extinction (dont get anything)

low because it was devalued before stimulus was devalued

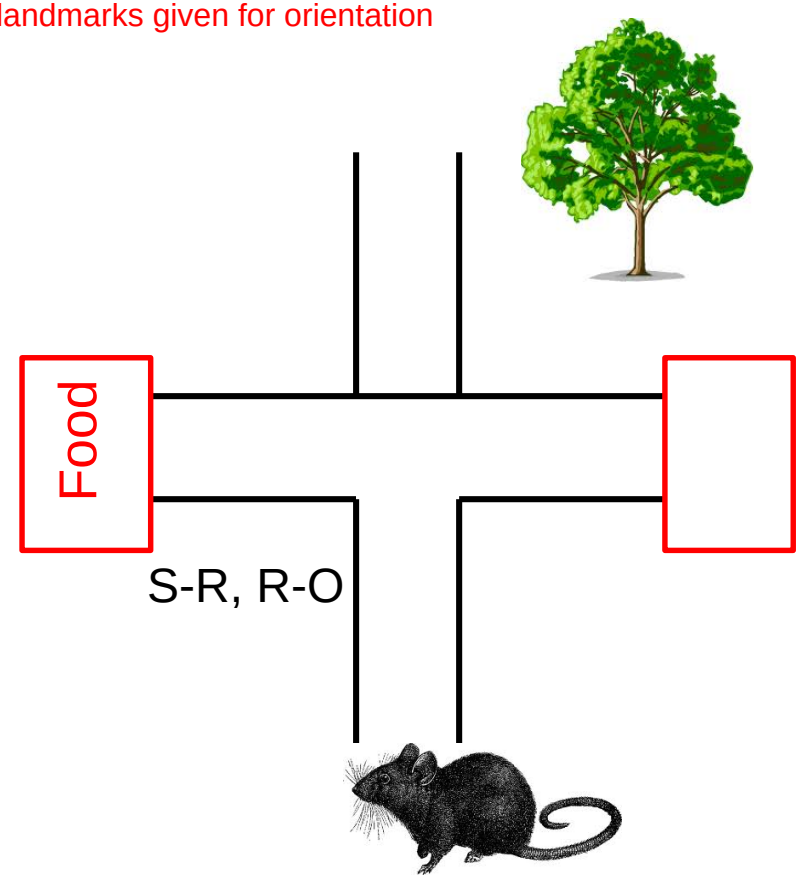
but law of effect was still learnt:

S=>R (habit, they do it without even thinking they dont think it will make them sick afterwards, they mindlessly do it so to say)

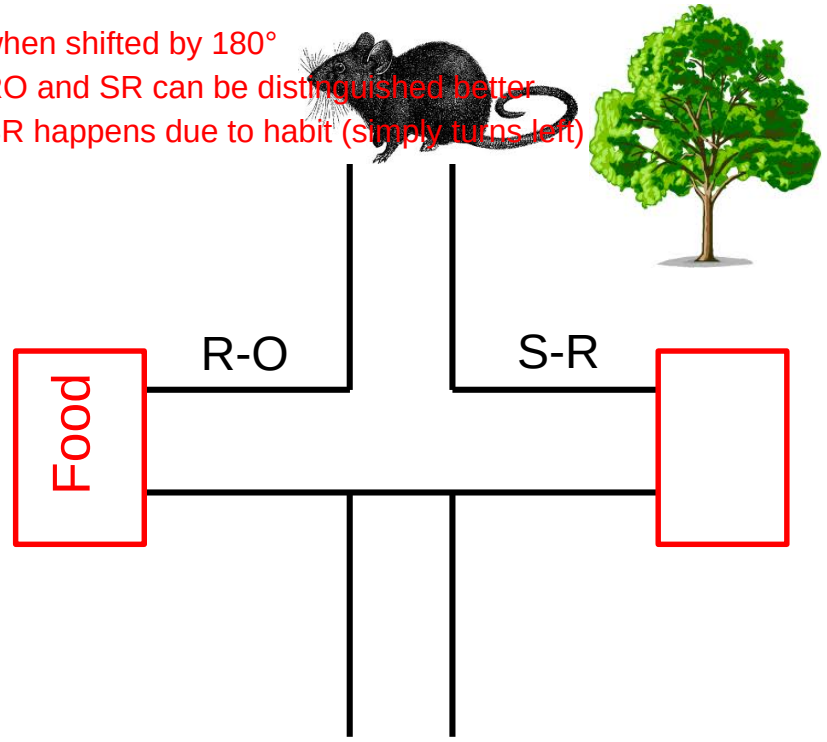
* R1 responding at low level even though food was now aversive, suggests S-R association

Response-UCS (expectancy, outcome) association and goal-directed behaviour

landmarks given for orientation



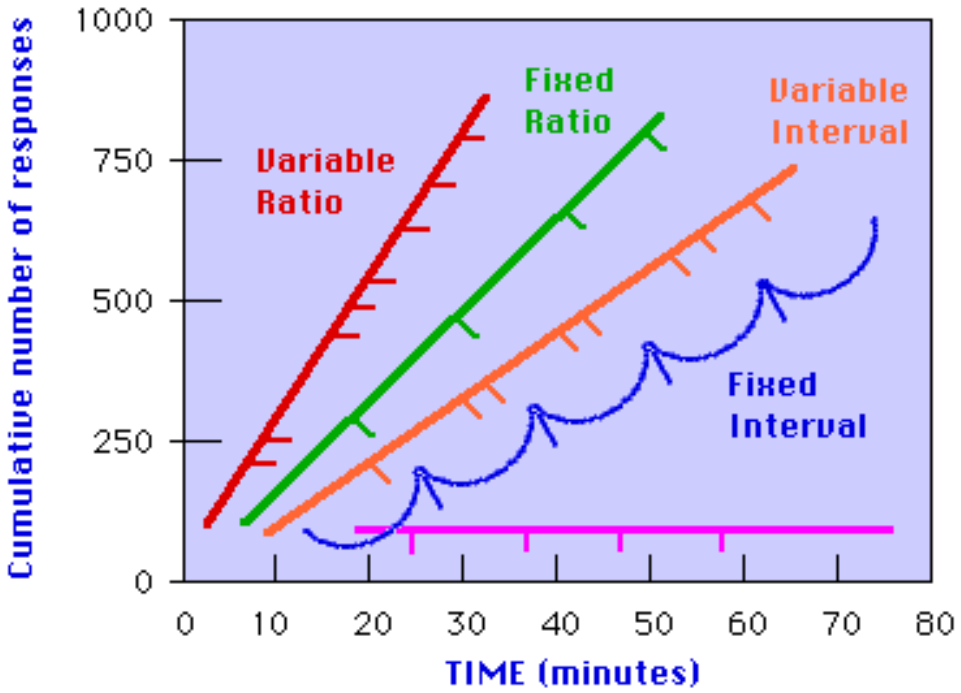
when shifted by 180°
RO and SR can be distinguished better
SR happens due to habit (simply turns left)



assumption: no concept of self, but there is still goal oriented behaviour
(tree there, food maybe there - no self-awareness with tree relative to I)

- “Walk forwards and turn left” (Stimulus-Response, S-R)
- “Use tree for orientation to get to goal” (Response-Outcome, R-O)

Schedules of reinforcement for operant conditioning



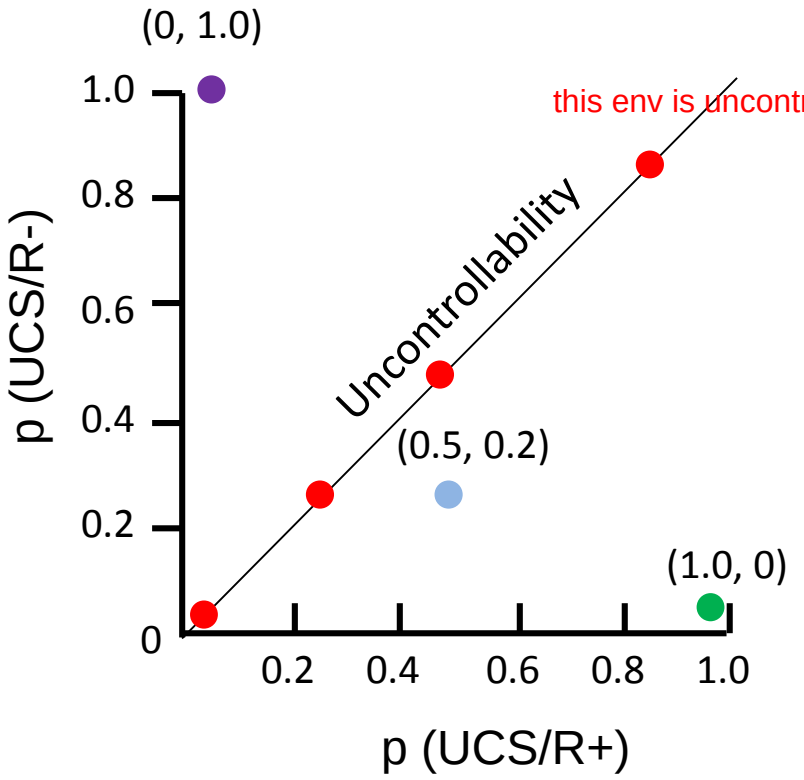
• Fixed-ratio: reinforces every n^{th} response that the subject makes since previous reinforcer
• Variable-ratio: reinforces the subject for emitting a particular number of responses, where the number of responses varies for each successive reinforcer

• Fixed-interval: reinforces the 1^{st} response that occurs after a specified period of time has elapsed since previous reinforcer
• Variable-interval: reinforces the 1^{st} response that occurs after a specified period of time has elapsed since previous reinforcer, where the time varies from interval to interval

this ratio is used when using drugs (cocaine, heroine) on animals

• Progressive-ratio schedule: reinforces the subject for emitting a particular number of reponses between reinforcers, where the number of responses increases for each successive reinforcer

The response-outcome contingency space



p (UCS/R+) = conditional probability of an outcome (UCS) following a response (R+)

p (UCS/R-) = conditional probability of an outcome (UCS) occurring in the absence of that response (R-)

UCS = Food, Stop Electro-shock

Subject always receives UCS for response R+ (p = 1.0), never receives UCS if no response R- (p = 0.0) (Continuous reinforcement)

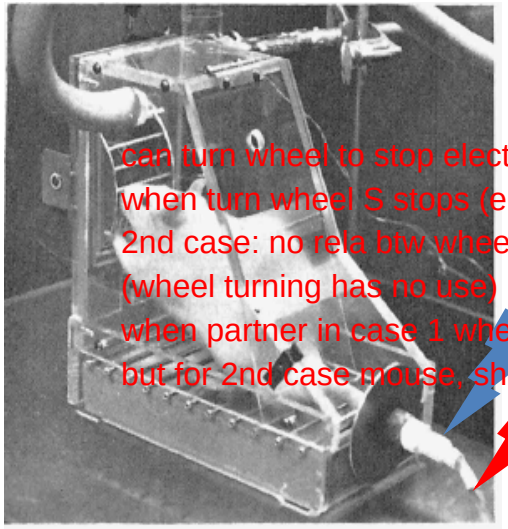
Subject receives UCS for R+ at p = 0.5 and receives UCS for R- at p = 0.2

Subject never receives UCS for R+, and always receives UCS for R-
(Differential reinforcement of other behaviour) find another behaviour

Subject receives UCS for R+, at the same rate as it receives UCS for R- (Uncontrollability)

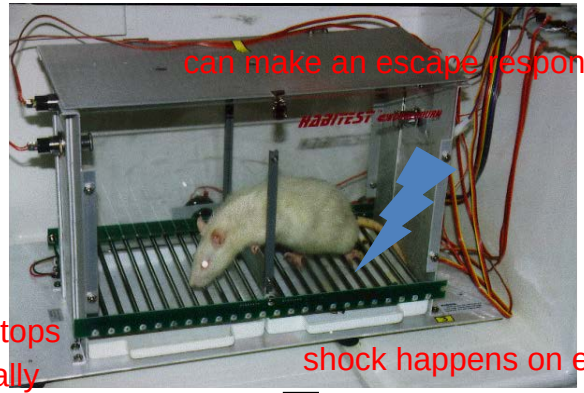
The learned helplessness (uncontrollability) effect in rats

Day 1
Electro-shock pre-exposure



can turn wheel to stop electro shock
when turn wheel S stops (elec shock)
2nd case: no rela btw wheel turning and elec shock stopping
(wheel turning has no use)
when partner in case 1 wheel turns and stops, his elec shock also stops
but for 2nd case mouse, shock stopping is completely random actually

Day 2
Escapable shock in 2-way Shuttle box



can make an escape response

shock happens on either side

TABLE 1

MEAN WHEEL-TURN ESCAPE LATENCIES ACROSS BLOCKS OF 10 TRIALS

Block	1	2	3	4	5	6	7	8
Latency (sec)	22.06	15.85	13.20	14.09	10.82	6.96	4.09	3.25

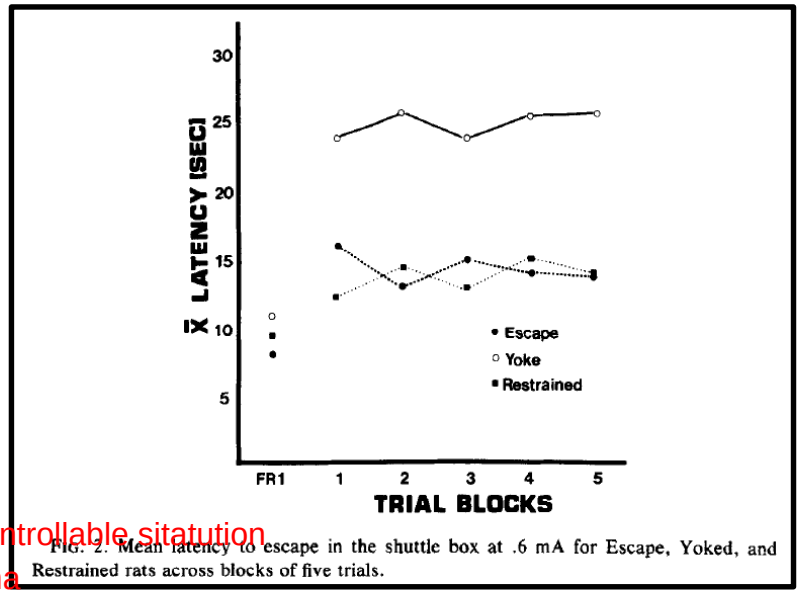
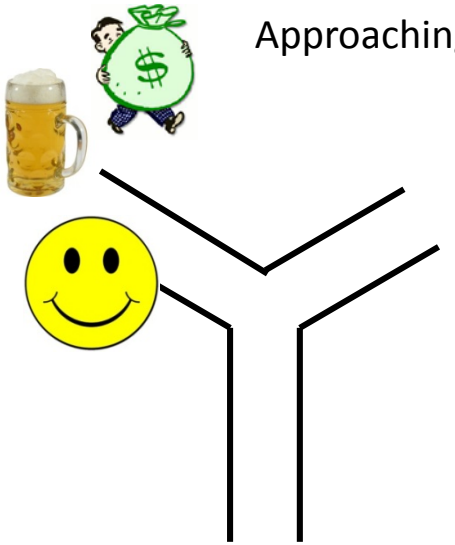


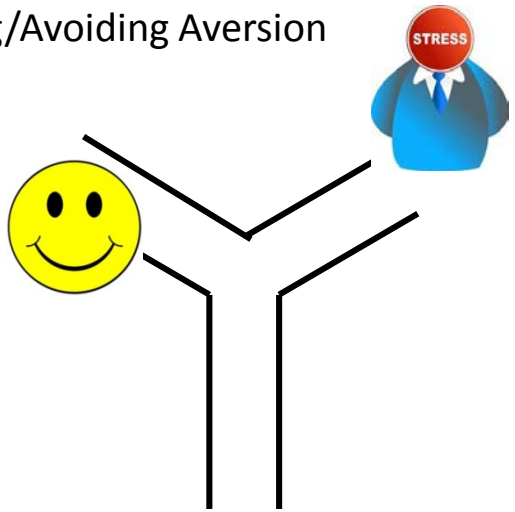
FIG. 2. Mean latency to escape in the shuttle box at .6 mA for Escape, Yoked, and Restrained rats across blocks of five trials.

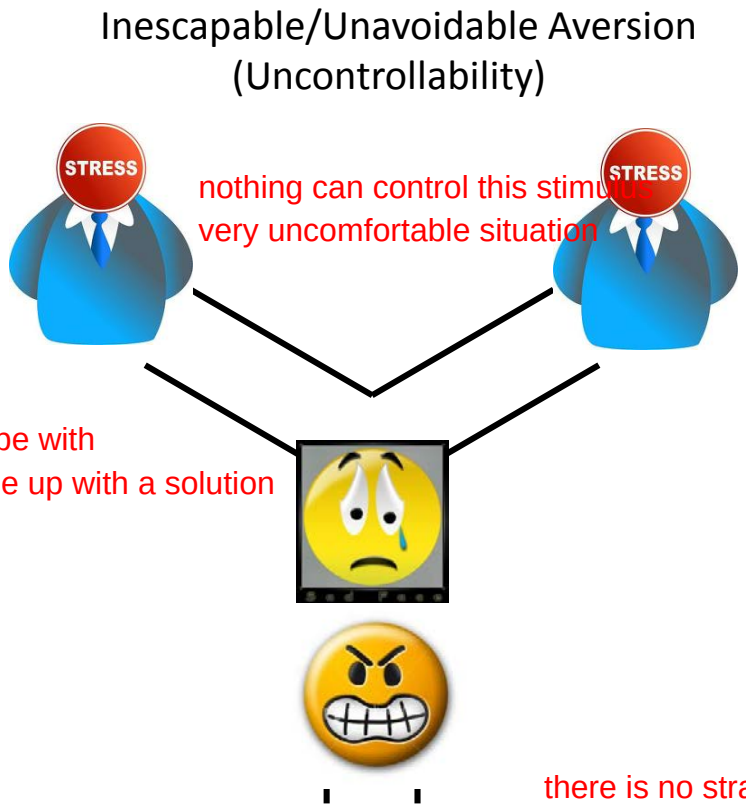
the prev case 1 mice escape quite fast
prev case 2 mice take twice as long because shock was an uncontrollable sitatution
if uncontrollablity has been learnt (helplessness), it will affect beha
in another behaviour
=> depression: specific operant uncontrollable exp happens in person and they suddenly treat every situation in their life as uncontrollable even though it is controllable

Jackson et al. (1978) Learn Motivation 9: 69

Individuals respond to and learn efficiently about environmental factors







it seems the brain never evolved to cope with uncontrollable situations - it never came up with a solution evolutionary

Contexts for Stressful life events:

- Employment
- Finance
- Health
- Housing
- Family
- Social relationships

there is no strategy to put right uncontrollable event, since event was uncontrollable (pissing off till even psychiatric condition)

Event (E1) → Event (E2) Association

Classical Conditioning:

ex: look for M-sign for migros => get food

E1 = neutral/conditioned stimulus → E2 = emotional unconditioned stimulus

these are two types of learning

Operant Conditioning:

behave nicely, so you get food

E1 = behaviour → E2 = emotional unconditioned stimulus

in foreign countries, operant conditioning is more prevalent, until one has found a supermarket to make an association with foodmarket sign

Integration of stimulus-stimulus and response-outcome conditioning

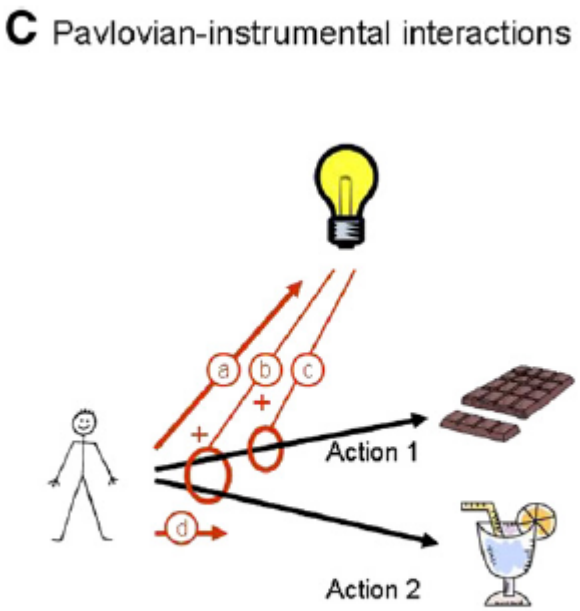
Conditioned reinforcement:

CS → UCS Response → CS → Responses

other word for operant

Pavlovian-instrumental transfer:

CS → UCS1 Response → UCS1 (or 2) CS + Response → Responses



a = Conditioned reinforcement
b = General Pavlovian-instrumental transfer
c = Specific Pavlovian-instrumental transfer
d = Approach responses/Pavlovian conditioned responses

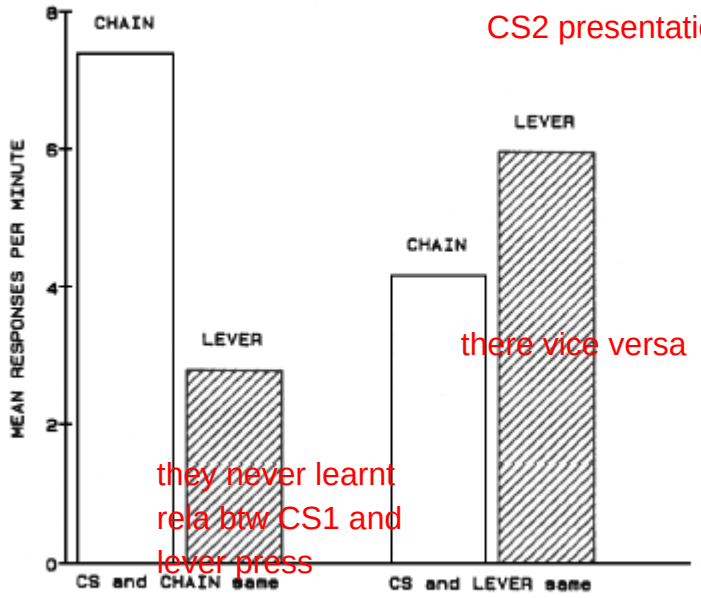
Pavlovian-to-Instrumental Transfer

CS1 → US Food
Chain pull → US Food
CS1 → Chain pull

CS2 → US Sucrose drink
Lever press → US Sucrose drink
CS2 → Lever press

CS1 presentation

CS2 presentation



there vice versa

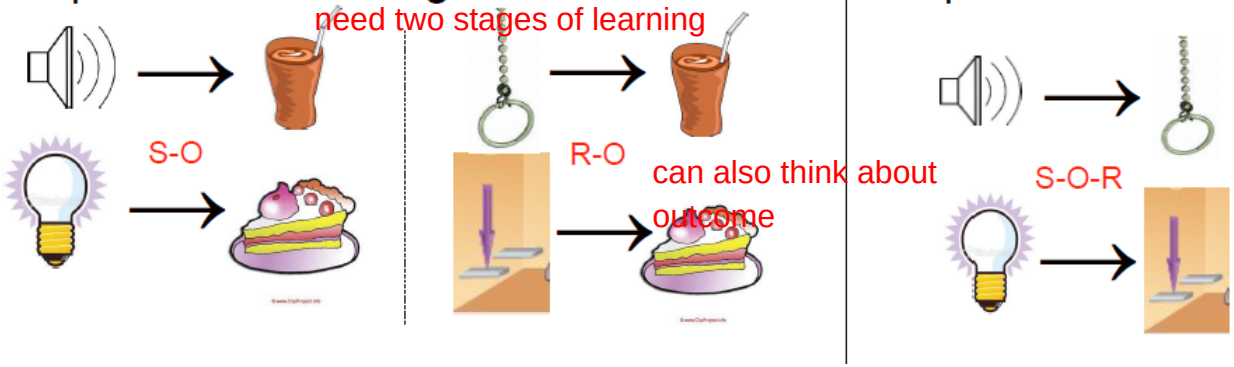
they never learnt rela btw CS1 and lever press

Note. Responding is shown during the presentation of a Pavlovian conditioned stimulus (CS) that signaled the same reinforcer as that earned either by a chain pull or a lever press.

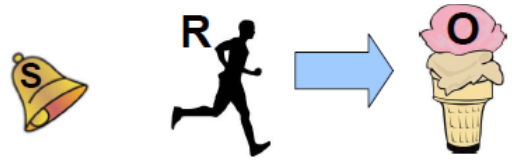
Associative theories of goal-directed behaviour

Outcome-Response theory (O-R)

- Pavlovian-to-Instrumental transfer (PIT)
- Pavlovian stimulus can prime performance of an instrumental response that was separately paired with the same outcome as the Pavlovian stimulus
- separate conditioning:



Response-Outcome theory (R-O)



- feed-forward model: decision process starts with the thought of the alternative courses of action available

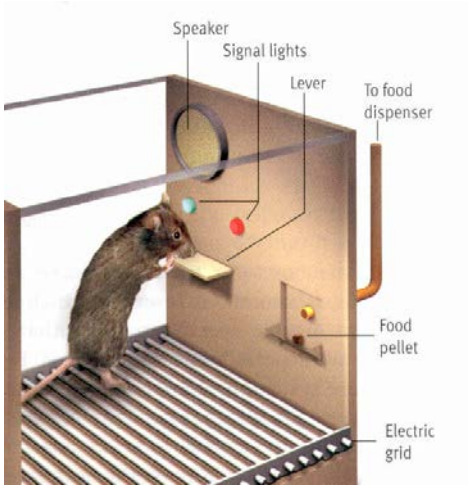
important slide

Association learning and Goal-directed behaviour

S = Stimulus
O = Outcome
R = Response

- **S-O**: Pavlovian conditioning; predictive relationship
- **S-R**: habit acquisition, context priming, behavioural autonomy
- **R-O**: representation of causal relationship, action selection
- **O-R**: backward association, action planning, response priming

- **S-R-O**: selection of goal-directed action
- **S-O-R**: PIT paradigm (indirect priming, S-O/O-R)



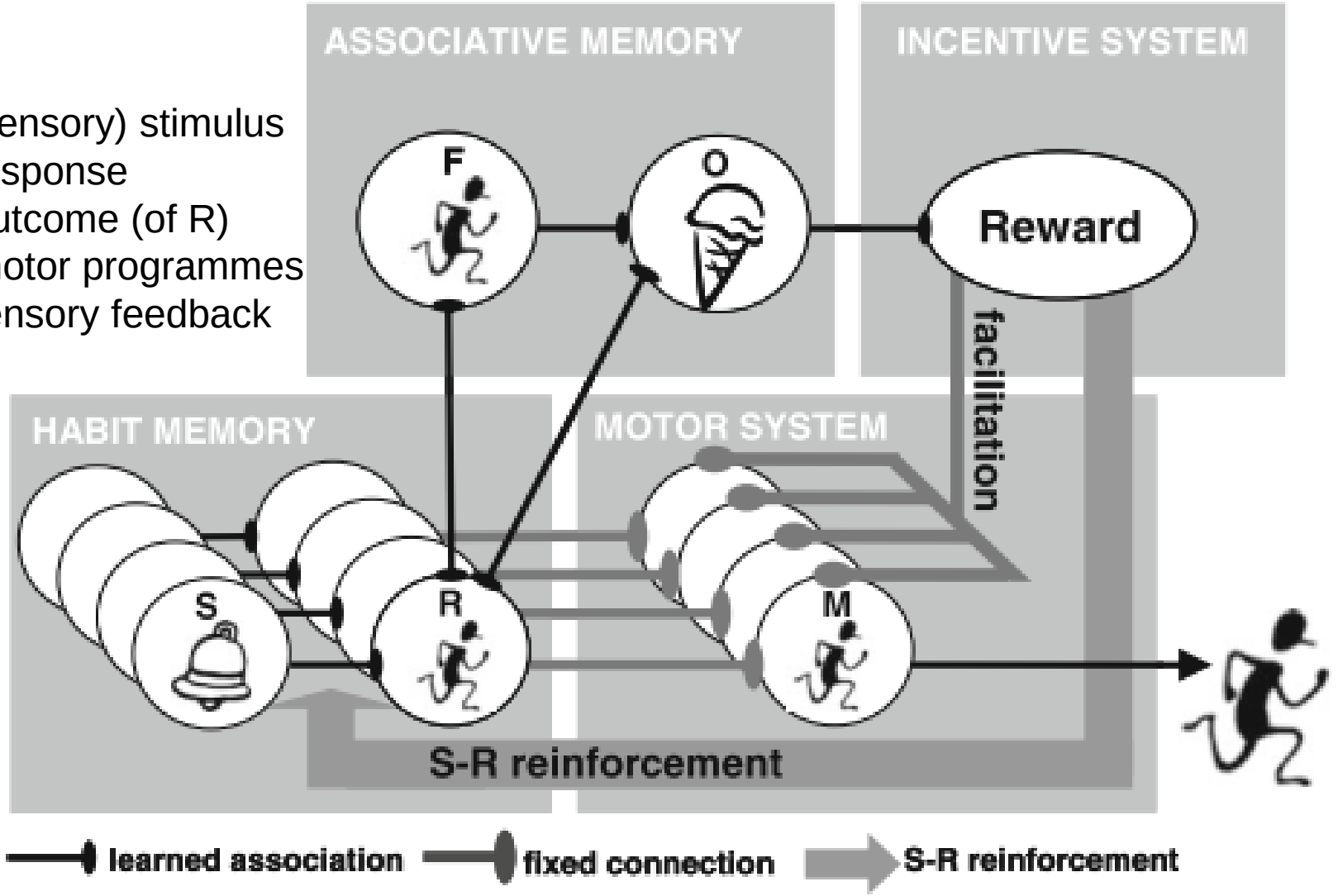
Criteria that need to be met for behaviour to be Goal-directed:
Belief criterion: There is a causal relationship between R and O
Desire criterion: O is an incentive (a goal)

De Wit & Dickinson Psychological Research (2009) 73: 463

see the paper of the second week that goes into depth of this model

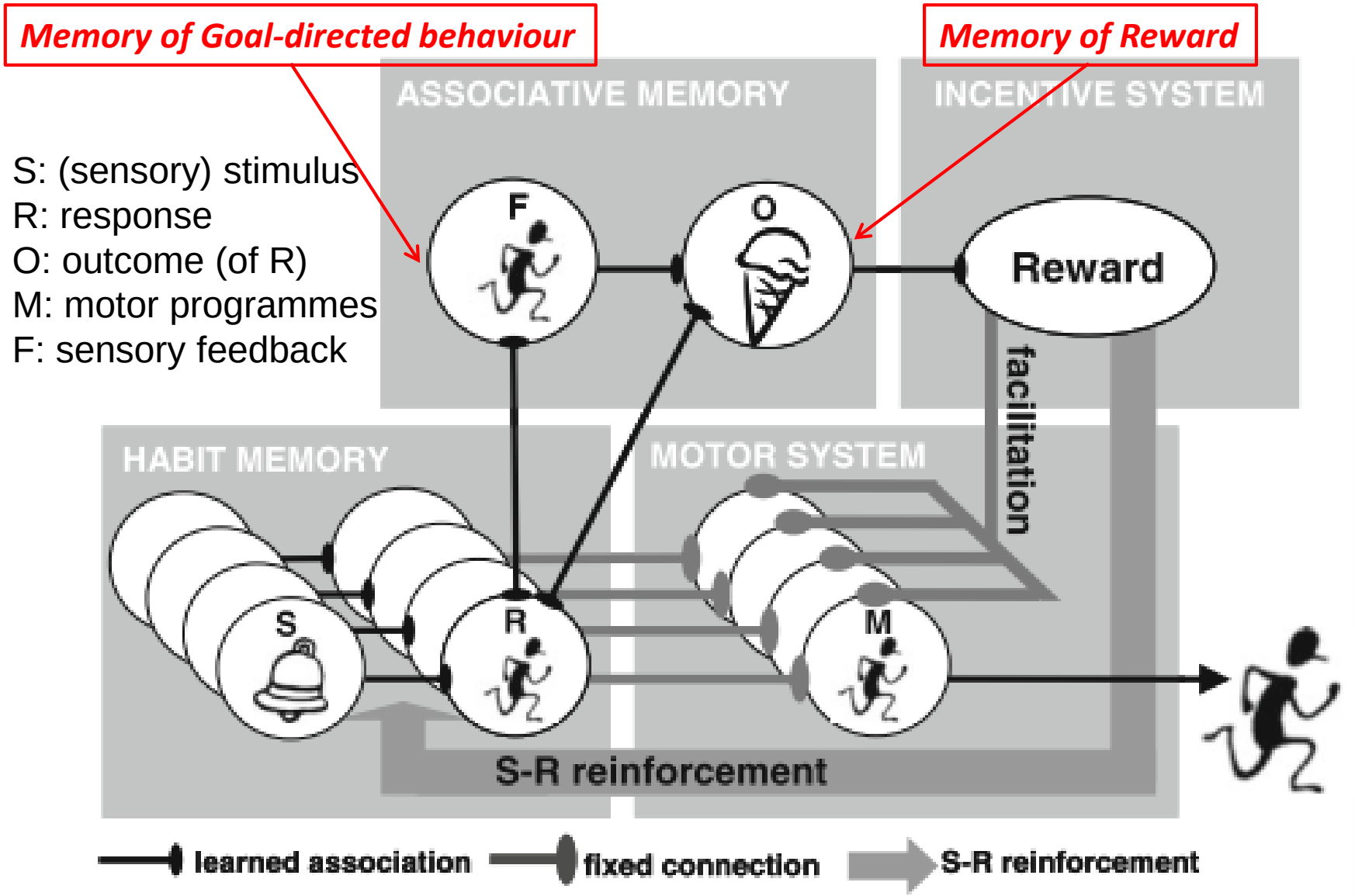
Associative-Cybernetic model

S: (sensory) stimulus
R: response
O: outcome (of R)
M: motor programmes
F: sensory feedback



De Wit & Dickinson Psychological Research (2009) 73: 463

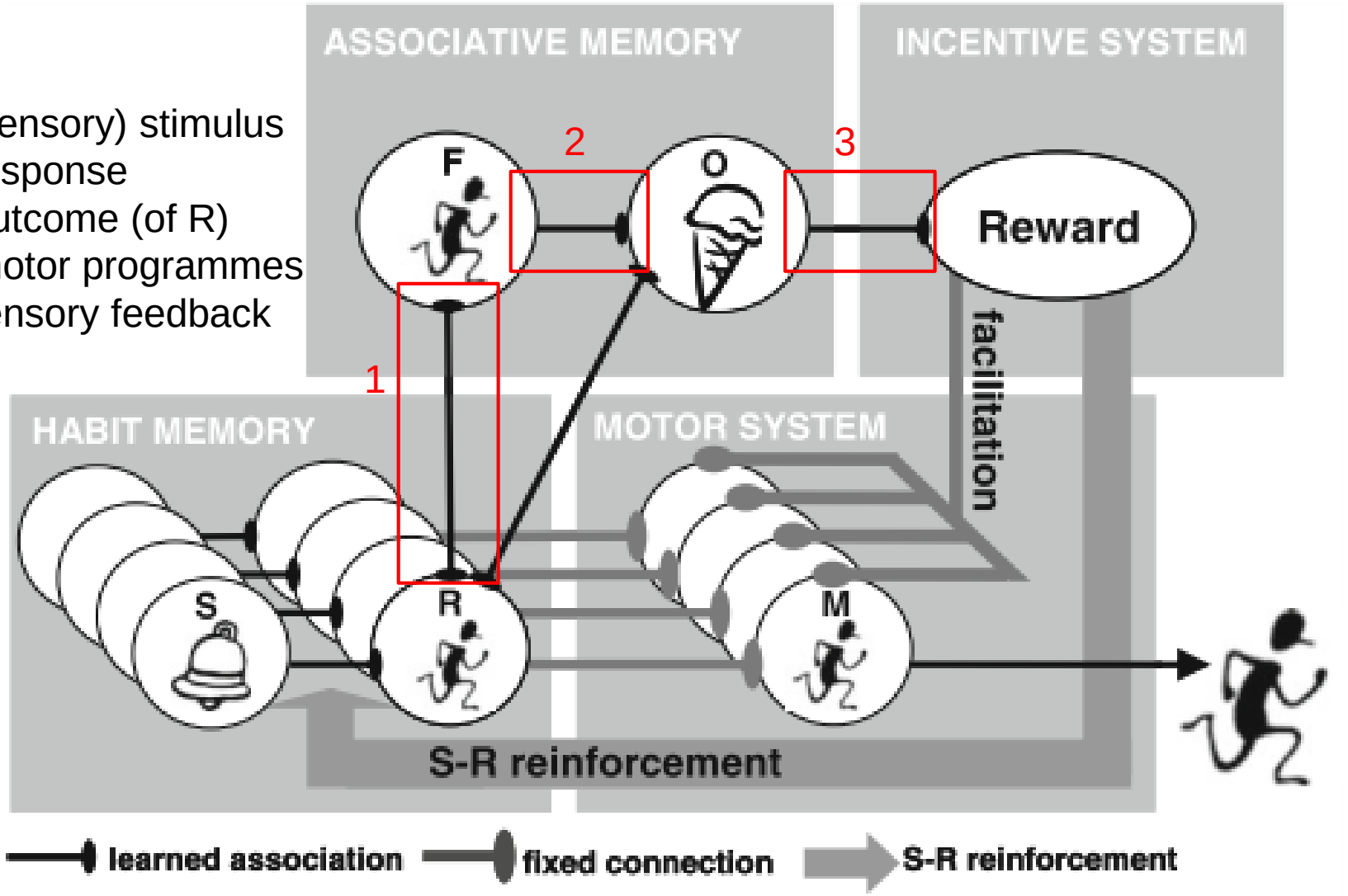
Associative-Cybernetic model



De Wit & Dickinson Psychological Research (2009) 73: 463

Associative-Cybernetic model

S: (sensory) stimulus
R: response
O: outcome (of R)
M: motor programmes
F: sensory feedback

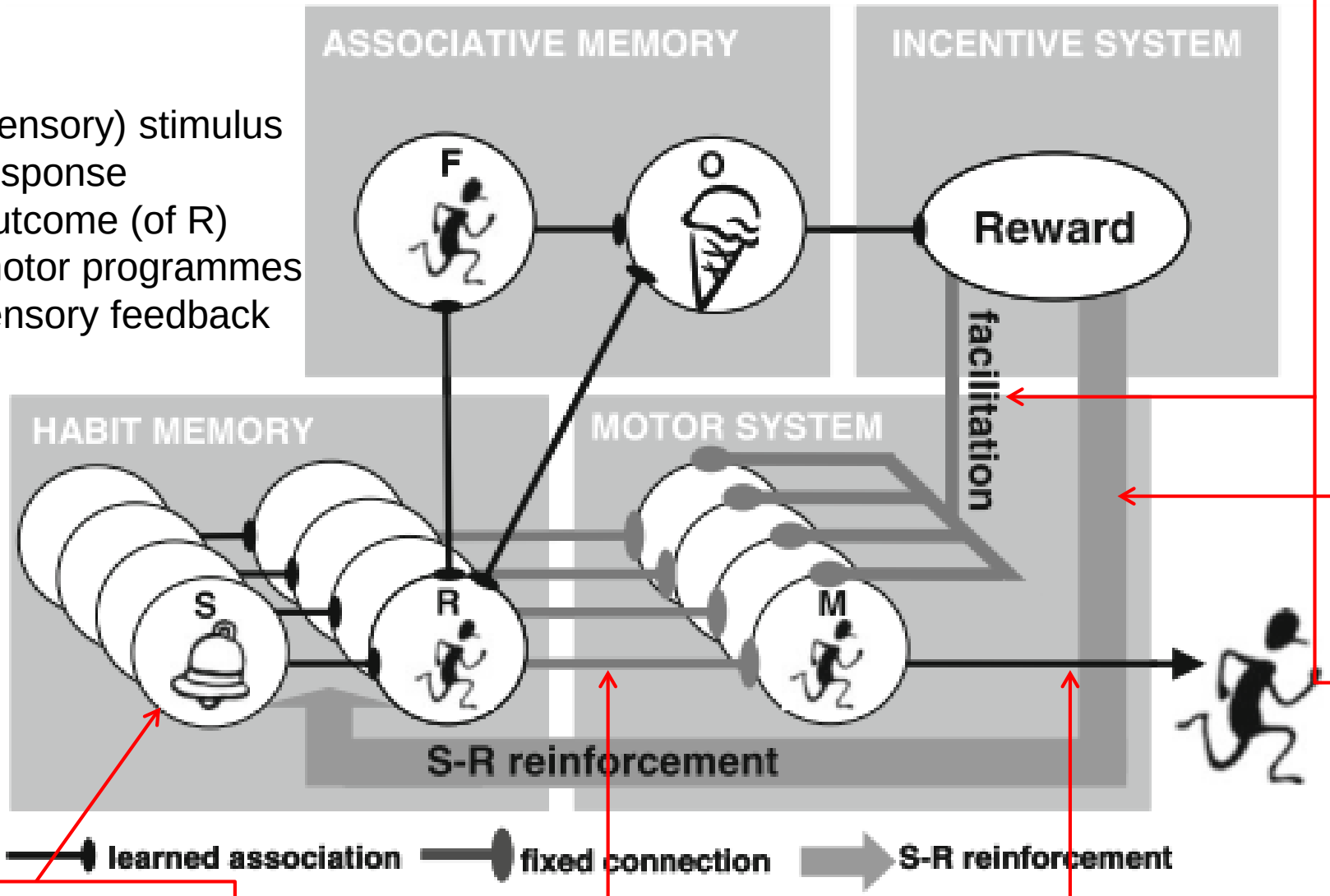


1. Activation of Response "Approach ice cream car" activates the thought of this Behaviour (F) (Goal)
2. Thought of approaching ice cream car activates thought of ice cream
3. Connection between sensory representation of ice cream (O) and its reward properties (Desire)

De Wit & Dickinson Psychological Research (2009) 73: 463

Associative-Cybernetic model

S: (sensory) stimulus
R: response
O: outcome (of R)
M: motor programmes
F: sensory feedback



Activation of Reward activates Motor Response

Reinforcement signal to Habit memory

Conditioned Stimuli activate Habit Response

Output of the Habit Memory

Motor programmes driving behaviour

De Wit & Dickinson Psychological Research (2009) 73: 463

Operant Conditioning and Goal-directed behaviour

• Animals can acquire behaviours that result in obtaining a reward or escaping an aversive stimulus

• Animals do not learn an association between an operant stimulus (e.g. lever) and a response, although such stimulus-response habits do contribute to operant behaviour

• Animals learn an association between their behaviour and the reward or aversive stimulus, with the reward/aversive stimulus being the expected outcome or goal of their behaviour

• Conditioned stimulus (CS-Outcome) can increase operant behaviours that lead to the same unconditioned stimulus (Response-Outcome), even though the classical conditioning and operant conditioning occurred separately. This CS-Outcome-Response or Pavlovian-to-instrumental transfer provides some of the best evidence for outcome/goal-directed behaviour

• In most situations, the CS energizes the response that leads to the outcome (CS-Response-Outcome)

• Reinforcement schedules describe the relationship between responses and outcome, and can vary in terms of number of responses required or time elapsed before a response will be reinforced

• Under conditions where there is no association between behaviour and outcome, the outcome is uncontrollable. Uncontrollable aversive stimuli lead to helplessness.