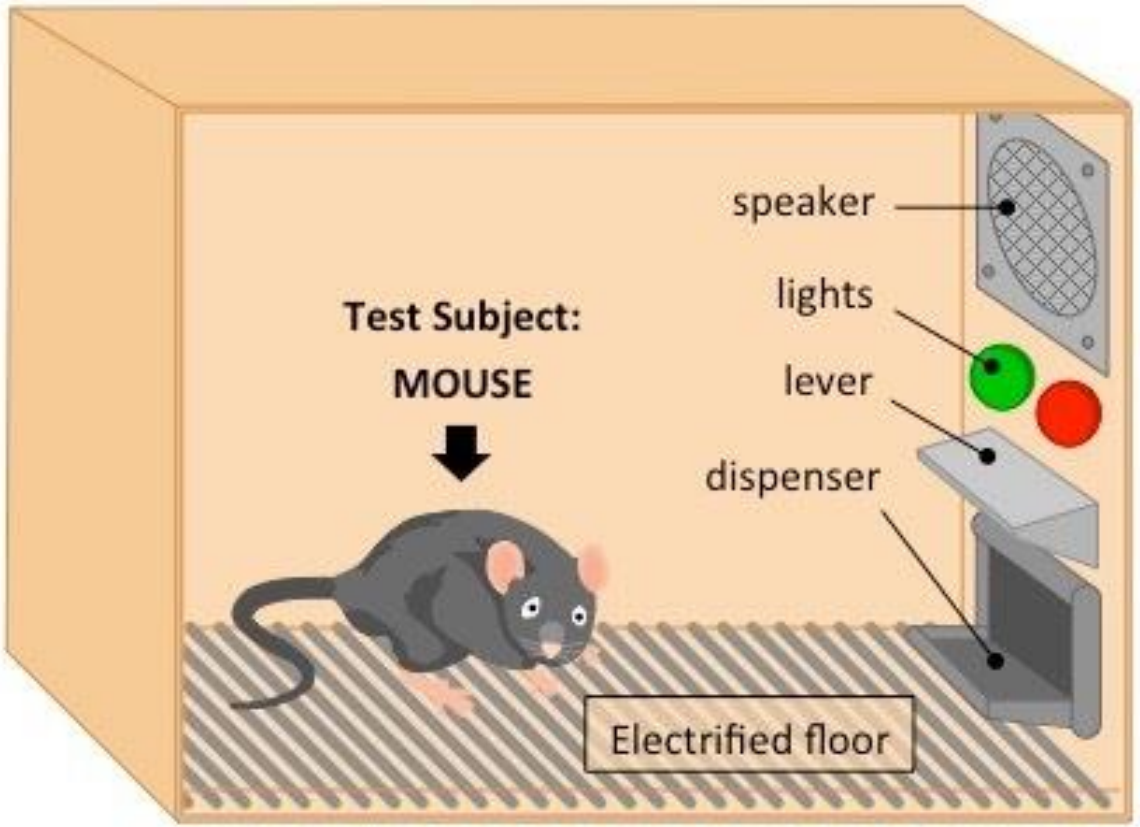
Predicting reward

- Getting a challan/ticket for speeding

vs

- License suspension for 1 year for DUI of alcohol.

Skinner’s experiments

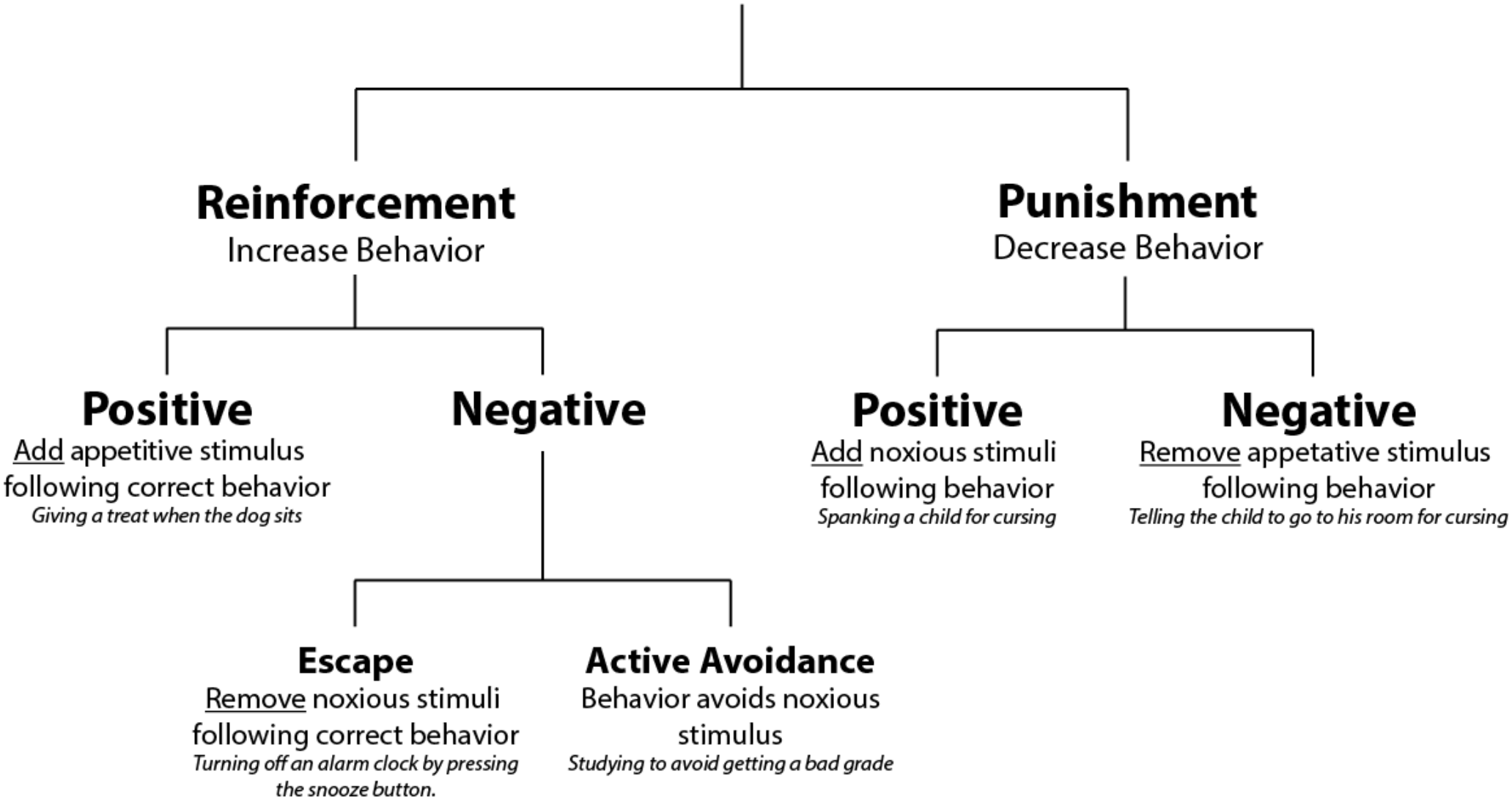


	Something **given** to the mouse	Something **taken** from the mouse
Increases likelihood of repeated behavior	**POSITIVE REINFORCEMENT** Mouse given food when lever pressed (after green light)	**NEGATIVE REINFORCEMENT** Loud noise stopped when lever pressed
Decreases likelihood of repeated behavior	**POSITIVE PUNISHMENT** Mouse is shocked when lever pressed (after red light)	**NEGATIVE PUNISHMENT** Not applicable in this scenario

Operant Conditioning Paradigms

	Something is given (positive)	**Something is taken away (negative)**
Increases response or behaviour (reinforcement)	Positive reinforcement Example: Clean room → get appreciation	Negative reinforcement (escape/avoidance training) Example (escape): Take aspirin → headache goes away Example (avoidance): Studying hard to avoid failing in exams
Decreases response or behaviour (punishment)	Positive punishment Example: (penalty/fine) → reduce driving speed	Negative punishment (omission training) Example: Fight with other children → time-out from play License suspension for 1 year for DUI of alcohol.

Operant Conditioning



At the grocery store, 2-year-old sees chocolates and wants it. Her mother says no, and the girl throws a tantrum. To calm her down, the mother relents and buys girl chocolates. The next time they go shopping, the girl sees chocolates and immediately throws another tantrum. This time, she gets the chocolates quickly.

Girl learns to throw tantrums to obtain chocolate. This is positive reinforcement.

- Scenario 1 is presented from the girl's point of view. But consider the same story from the mother's point of view:

Mother takes her daughter to a grocery store. The child sees chocolate, wants it, and throws a tantrum. Overtired and in a rush, the mother gives the child chocolate, and the tantrum stops. On the next trip, as soon as the child starts to cry, the mother quickly hands over some chocolate to stop the screaming.

Mother learns to give chocolate to stop the tantrums. This is negative reinforcement.

Kumar's football team has a no-alcohol policy: players sign pledges not to drink alcohol during the football season. One night, Kumar goes out with some friends and has a few beers. The coach finds out and revokes Kumar's playing privileges for a week. When allowed to rejoin the team, Kumar is careful to stay away from alcohol for the rest of the season.

Kumar learns not to drink alcohol during football season. This is negative punishment.

The coach is decreasing the drinking behavior (punishment) by taking away privileges/suspension (negative) .

Mrs. Qureshi installs an electric fence system around the perimeter of her yard and gives her dog a collar that makes a high-pitched noise whenever he gets too close to the boundary. The first time the dog strays out of bounds while wearing the collar, the noise plays and distresses him. Soon, the dog learns to avoid the noise by staying inside the yard.

Positive punishment transformed into Negative punishment

Raghav is a 10-year-old who hates PT class. One day, after eating the school lunch, he gets a stomachache. He tells his teacher that he is feeling sick, and the teacher allows him to skip PT class that afternoon. Now, on days when there is PT class, Raghav frequently feels sick after eating lunch at school.

Raghav learns to feel (or report) sickness in order to avoid PT class. This is negative reinforcement.

Reinforcement Need Not Follow Every Response

- An experimenter can also control the *frequency* with which outcomes are delivered

- **Continuous reinforcement schedule:** a reinforcement schedule in which every instance of the response is followed by the reinforcer
- **Partial reinforcement schedule:** a reinforcement schedule in which only some instances of the response are followed by the reinforcer

- The rules determining how and when outcomes are delivered in an experiment are called **reinforcement schedules**.

Reinforcement Schedules

Reinforcement is delivered after a predictable number of responses

Effect: produces high rate of response

Reinforcement is delivered at predictable time intervals

Effect: responses increase close to the time for next reward

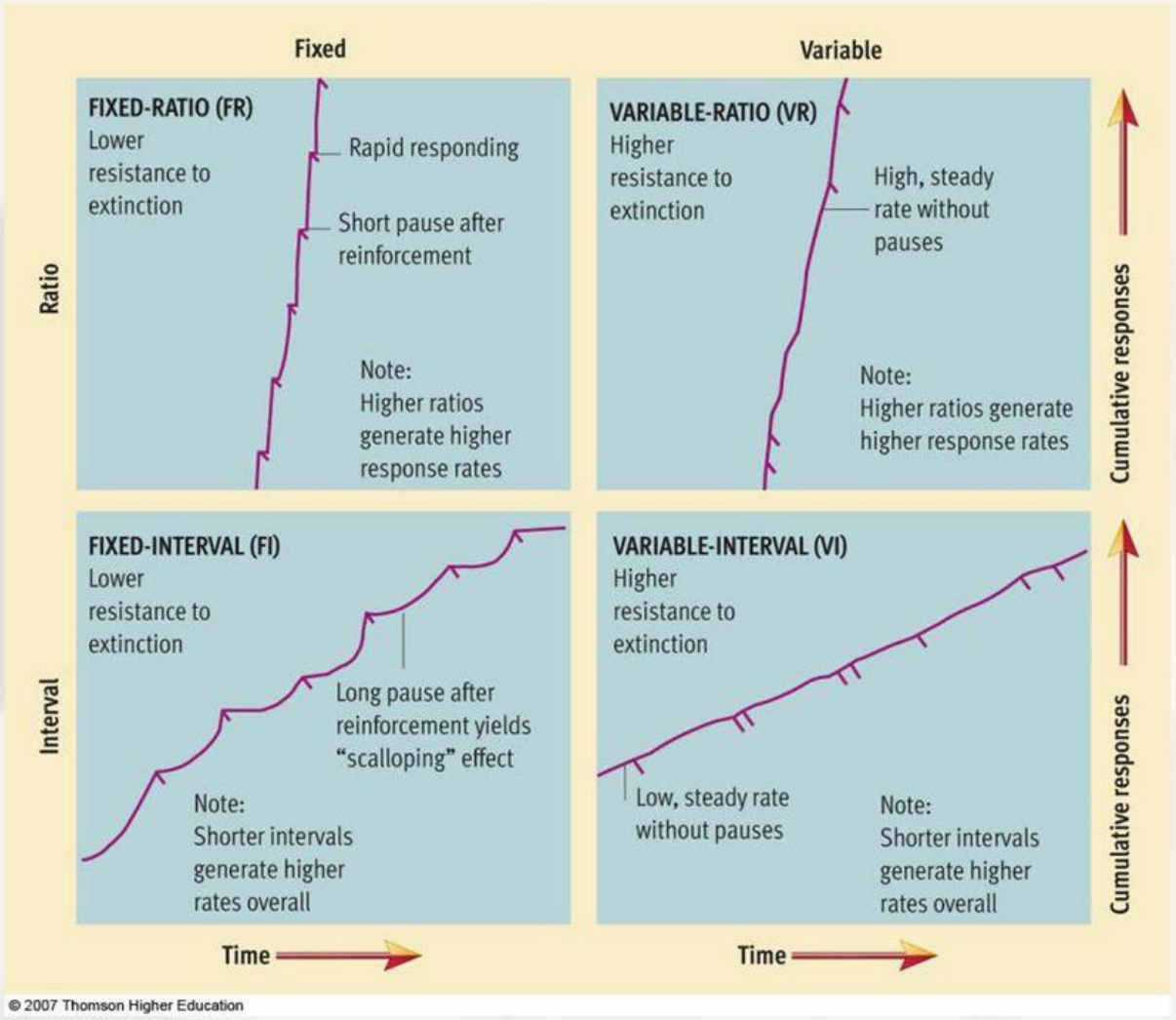


Figure 6.13 Schedules of reinforcement and patterns of response

Reinforcement is delivered after an unpredictable number of responses

Effect: produces a high steady rate of response

Reinforcement is delivered at unpredictable time intervals

Effect: produces a moderate, steady rate of response

Reinforcement Schedules

Gambling? VR

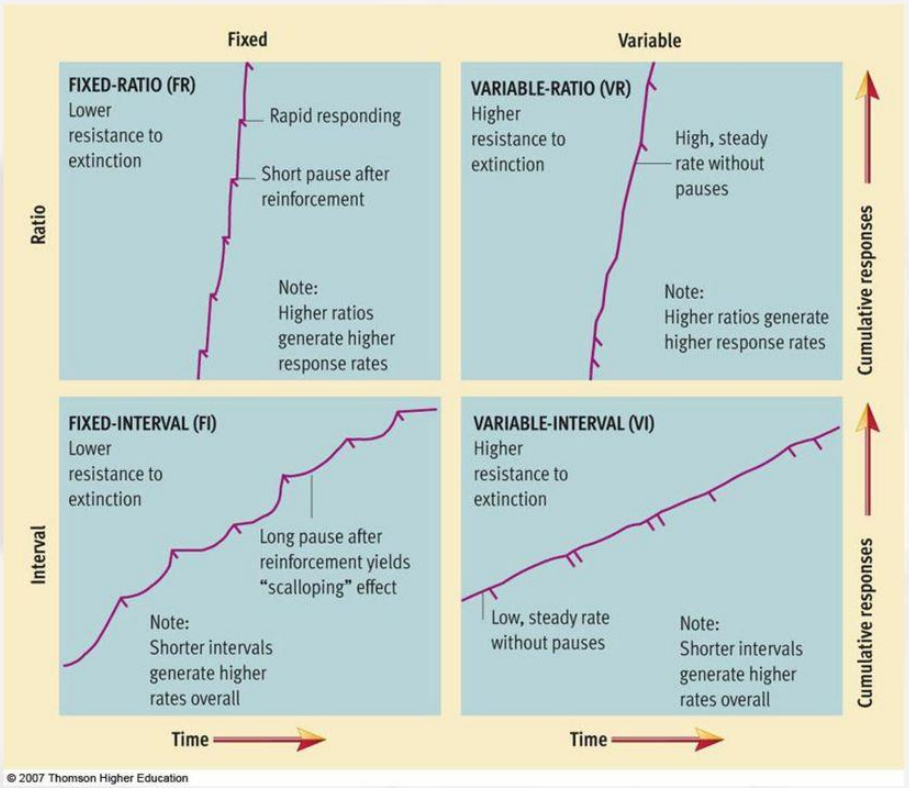


Figure 6.13 Schedules of reinforcement and patterns of response

Each first-grade student who completes the day's math worksheet receives a gold star from the teacher; at the end of the week, five gold stars can be exchanged for a toy. FR

A good telemarketer makes 20 calls and manages to sell his product to two customers, so he earns more profit if he makes a lot of calls. VR

Manish is a restaurant owner and has trained his employees to keep the restaurant clean with optimal service at all times. Every now and then the food inspector and quality control agents visit his restaurant. VI

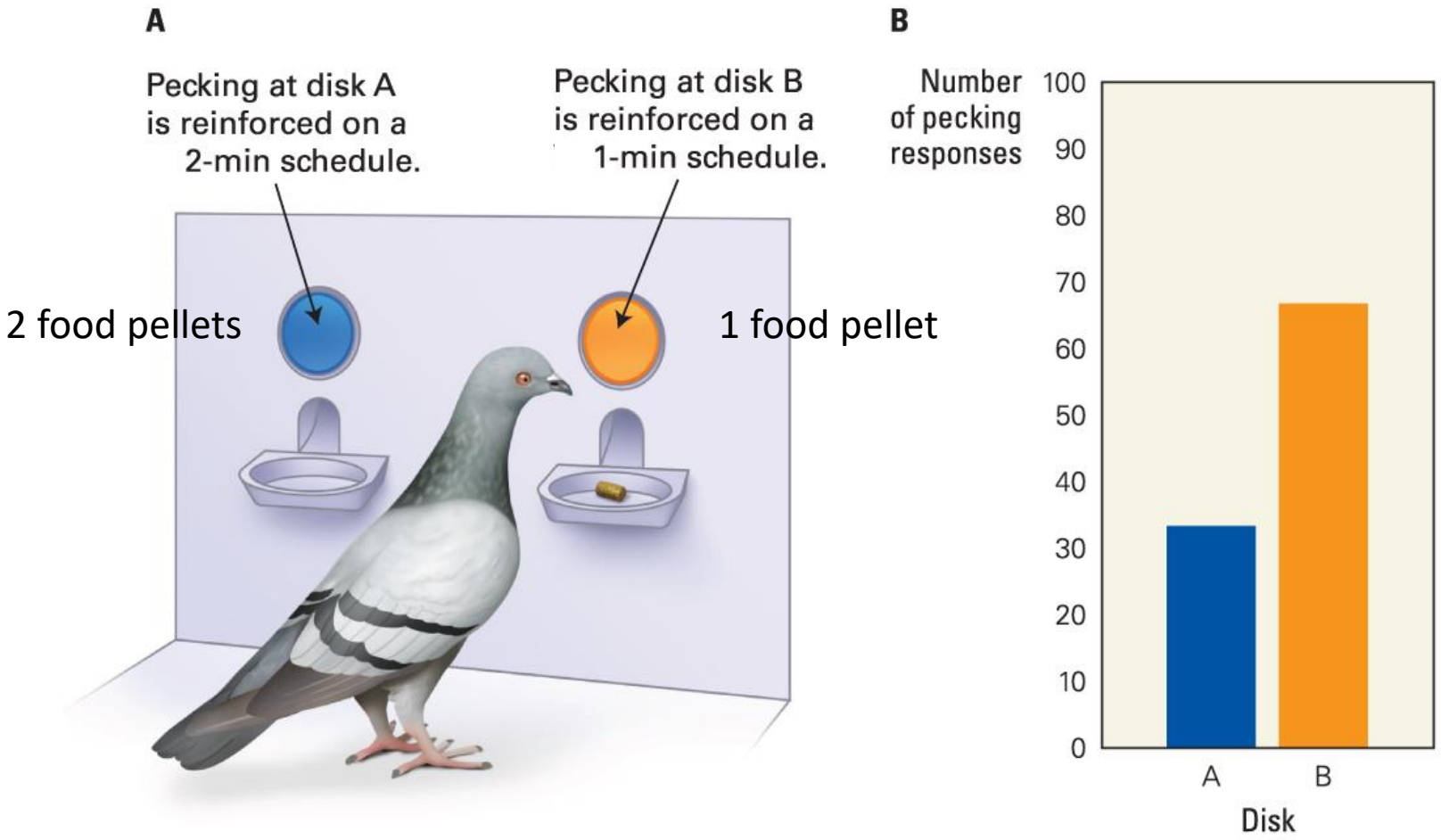
Maria donates blood regularly at the local hospital; they pay her for her donations, and it makes her feel good to know she's helping people in need. However, due to hospital policy, donors have to wait at least 2 weeks between donations. FI

A factory worker gets paid for every 10 clothes stitched FR

Getting random surprise quizzes in class. VI

Scheduled mid-term/end-term exams. FI

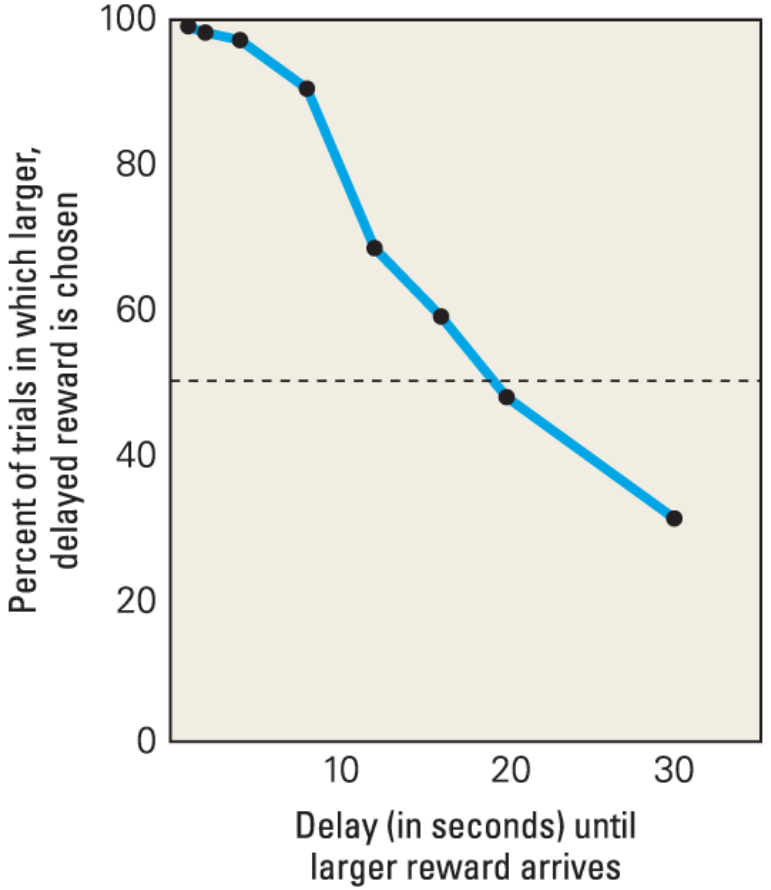
Choice behaviour



Gluck et al., *Learning and Memory*, 4e, © 2020 Worth Publishers

Delay Discounting

- **Delay discounting:** the progressive reduction (or discounting) of the subjective value of a reward the longer it is delayed
- For example, it is easy to put off having fun in order to study if the exam is coming up tomorrow; it is harder to wait if the exam is 5 weeks later, even though starting to study early will result in a better grade

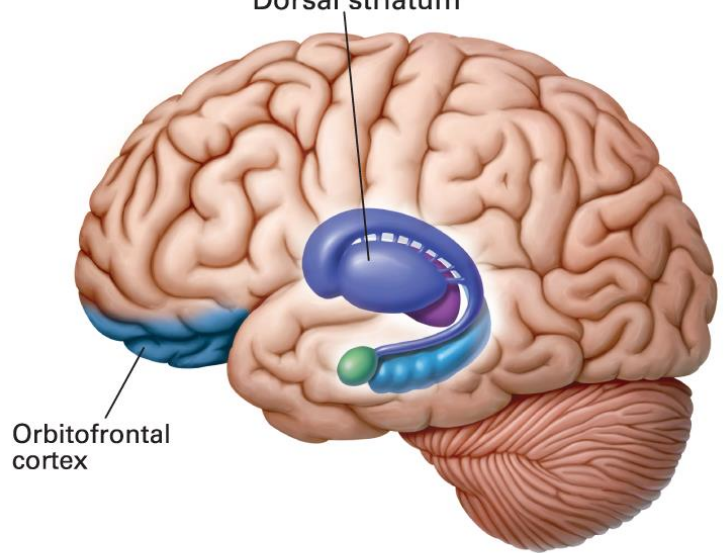


Gluck et al., *Learning and Memory*, 4e,
© 2020 Worth Publishers

Rats were given a choice between two levers: one that delivers a small, immediate reward (1 food pellet) and one that delivers a larger reward (3 pellets) after some delay

If that delay was short (e.g., 0 to 5 seconds), rats would choose the larger reward on nearly 100% of choice trials; but as the delay grew longer (e.g., 20 to 30 seconds), they were increasingly less likely to wait, opting more often for the smaller, immediate reward

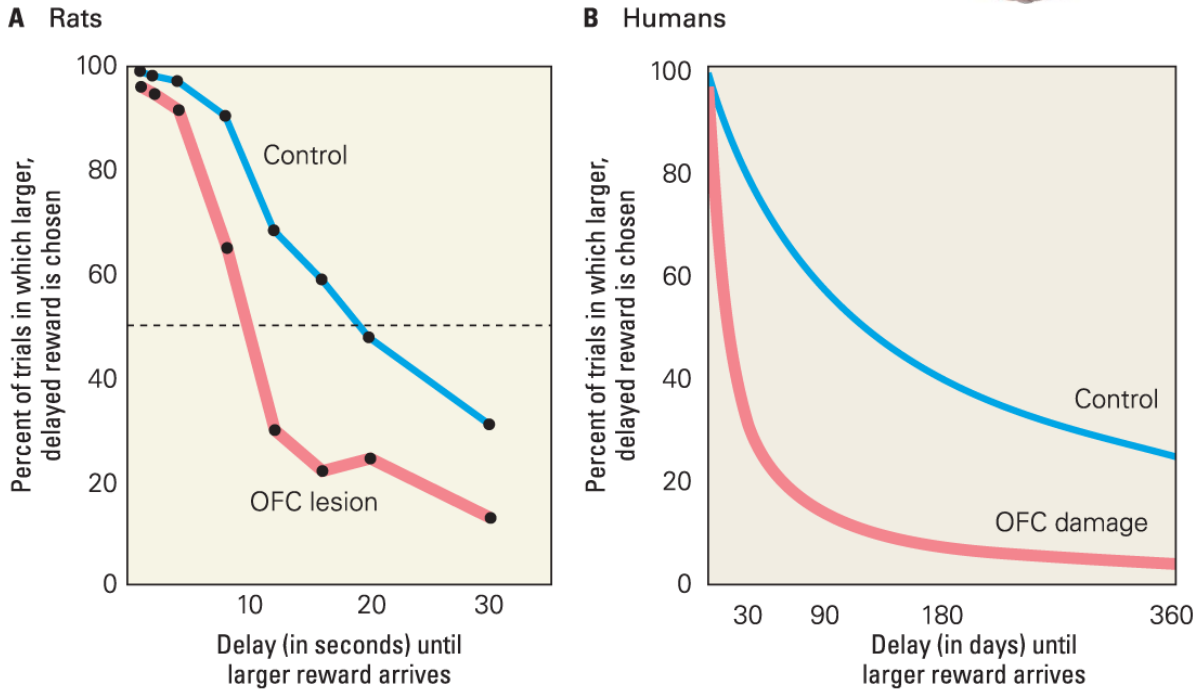
Delayed Reinforcement and the Orbitofrontal Cortex



Orbitofrontal cortex and self-control

(A) If control rats are given a choice between an immediate small reward and a larger, delayed reward, they will almost always prefer the larger reward if the delay is short (e.g., a few seconds), but as the delay increases, they will increasingly prefer the immediate reward. Rats with lesions of the orbitofrontal cortex (OFC) show much less self-control, indicated by a greatly decreased willingness to wait for the larger, delayed reward.

(B) Humans show a similar pattern: at short delays, both healthy (control) participants and people with OFC damage will usually opt for the larger reward; but as the delay increases, the people with OFC damage show much less self-control, indicated by decreased willingness to wait for the larger, delayed reward.



Gluck et al., *Learning and Memory*, 4e, © 2020 Worth Publishers

orbitofrontal cortex evaluates → expected rewards + expected timing of those rewards

Delayed Reinforcement (gratification) and Self-Control

- **Self-control** refers to an organism's willingness to forgo a small immediate reward in favor of a larger future reward
- One way of improving an individual's ability to wait for a reward is to induce the person to make a *precommitment*—that is, to make a choice that is difficult to change later

Delayed gratification

Taken from Iddo Landau *Finding Meaning in an Imperfect World*

Problem of Delayed Reward

We are too successful in delaying gratification, which means we cannot stop to enjoy any of our achievements. We cannot rest on our laurels and appreciate what we have. We must always delay reward to the future.

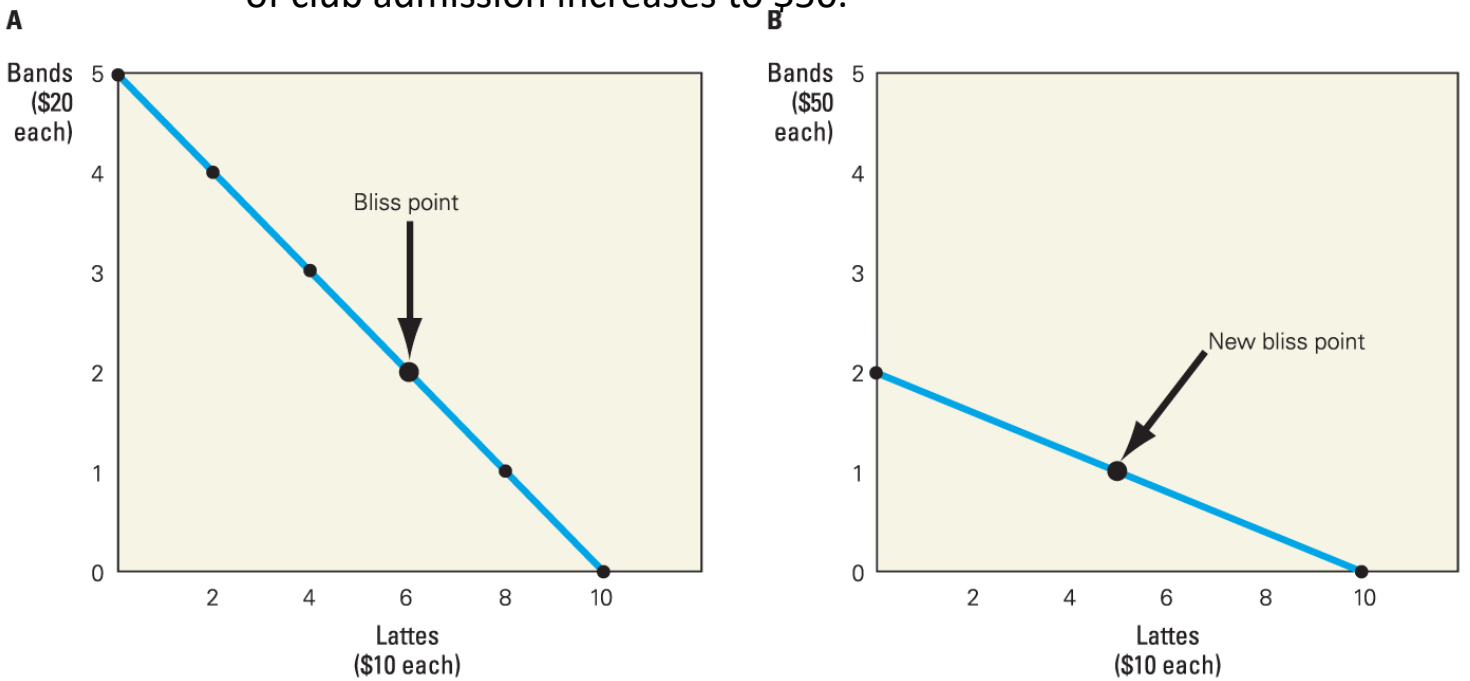
1. **Workaholism**
   We always have to be busy, to be doing stuff, to be productive.
2. **Stinginess in Compliments**
   We must always find fault in what we do and what others do. We must always find the defect that needs to be fixed.
3. **Hyper-competitiveness**
   We must always compare our achievements to those of others; we must always find someone who is better than us in some respect.
4. **Overselling the future**
   We expect too much from future rewards, so much so that when they arrive we are inevitably disappointed.

Behavioral Economics

- **Behavioral economics** is the study of how organisms allocate their time and resources among possible options
- Economic theory predicts that each consumer will allocate resources in a way that delivers the maximum "subjective value," or relative satisfaction, for that person
- The particular allocation of resources that provides maximal subjective value to an individual is called the **bliss point**

Optimizing gain/loss – no direct operant conditioning

(A) A student with $100 income per week may choose to distribute it between buying expensive coffee ($10 per cup) and seeing live music bands ($20 club admission); any point on the line will satisfy the budgetary constraints. The bliss point is the point at which this particular individual gets maximum subjective value for his money. (B) The bliss point may shift as conditions change—for example, if the cost of club admission increases to $50.



Gluck et al., *Learning and Memory*, 4e, © 2020 Worth Publishers

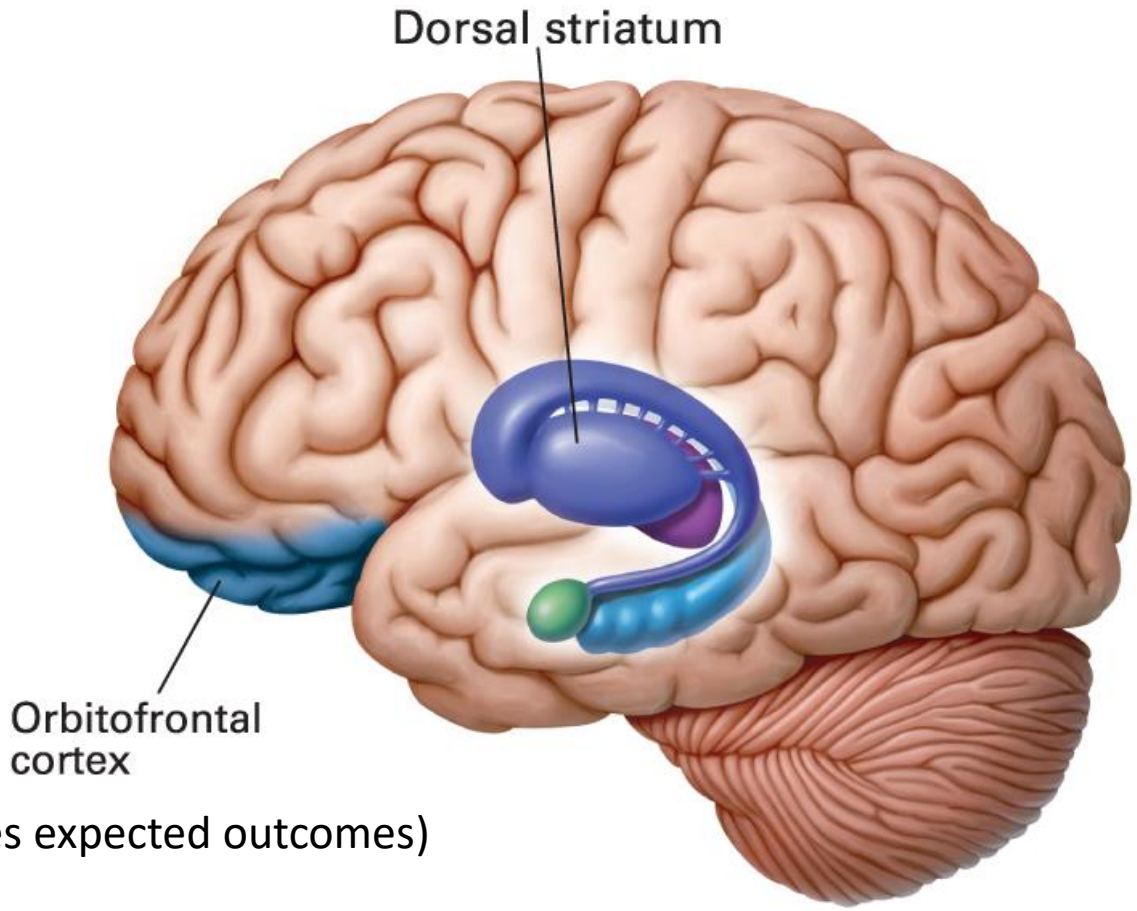
From a behavioral economics perspective, operant conditioning is not so much training an organism to execute a specific behavior as causing the organism to shift its allocation of time and energy among possible behaviors

Brain Substrates

- Two key areas
  - Dorsal striatum
  - Orbitofrontal cortex (expected outcomes)
    - Receives inputs that convey the full range of sensory info (visual, auditory, etc,) + visceral sensations (pain, nausea, etc)

(determines response and plans motor behaviour)

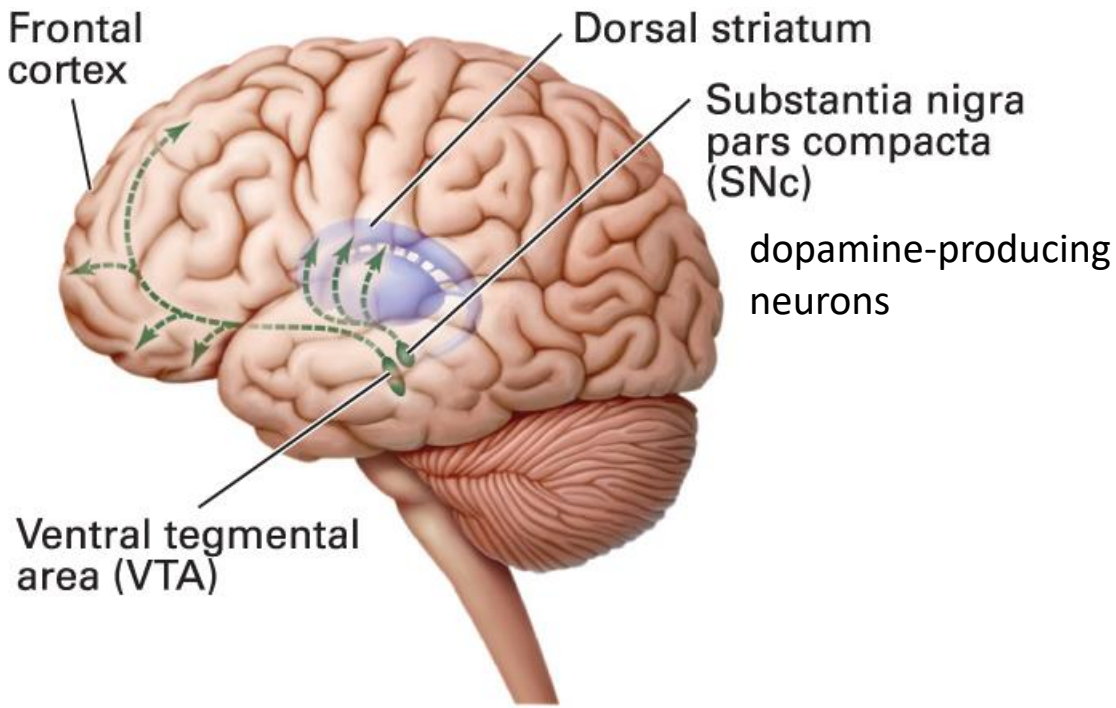
Part of the basal ganglia



(evaluates expected outcomes)

Gluck et al., *Learning and Memory*, 4e, © 2020 Worth Publishers

“Wanting” and “Liking” in the Brain



Gluck et al., *Learning and Memory*, 4e, © 2020 Worth Publishers

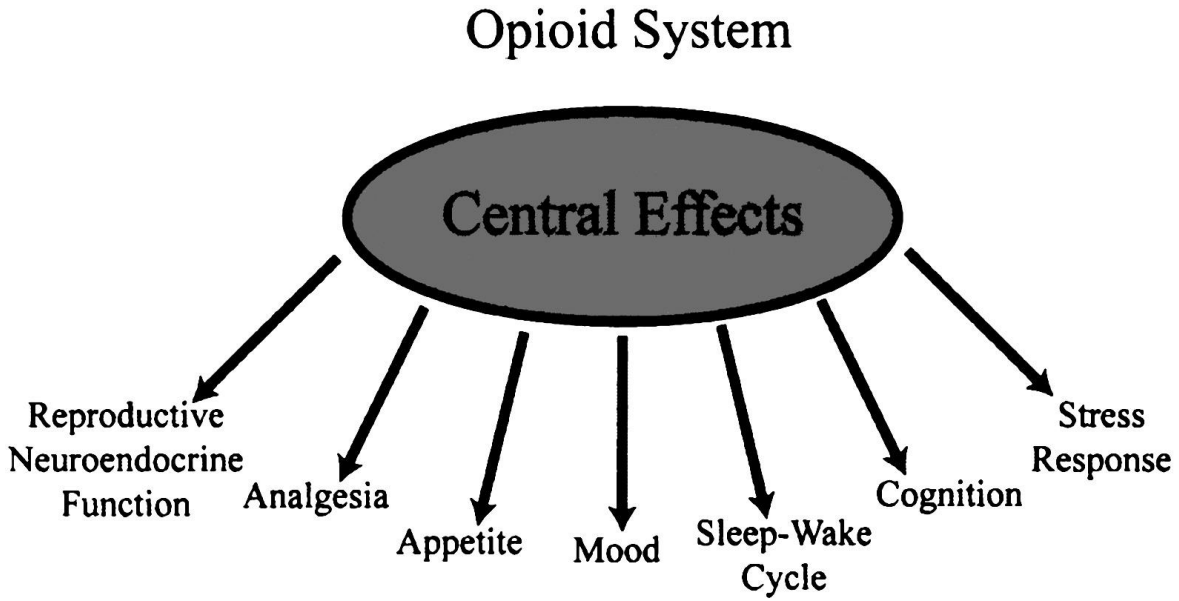
- Studies show that rats will work for electrical stimulation in several brain areas, in particular the **ventral tegmental area (VTA)**, a small region in the midbrain of rats, humans, and other mammals

- Many researchers believe that we have separate brain systems for signaling **hedonic value**—meaning the subjective “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 willing to work to obtain it.

- **incentive salience hypothesis** of dopamine function states that one role of dopamine in operant conditioning is to signal how much the animal “wants” a particular outcome—that is, how motivated the animal is to work for it

Endogenous Opioids Signal “Liking”

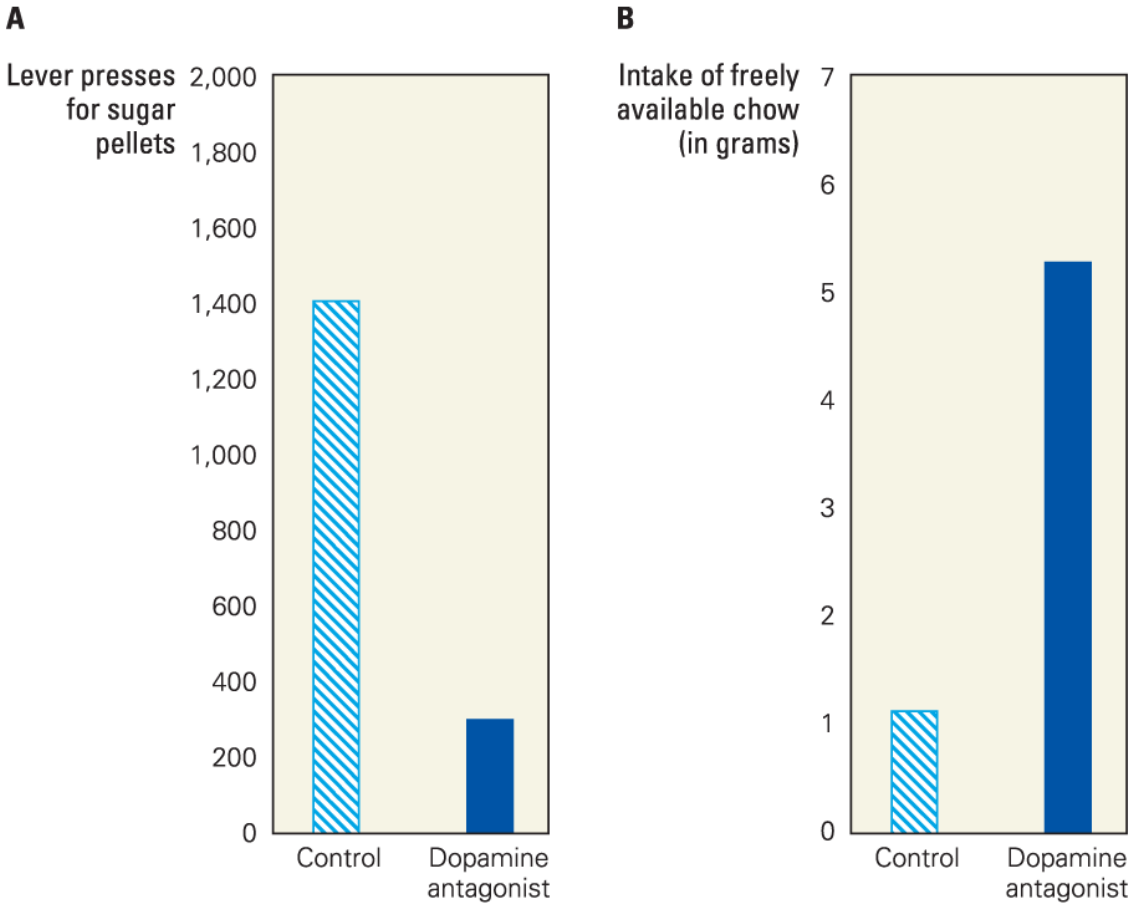
“liking” via the opioid system



- The **endogenous opioids (make you feel good, suppress pain)** are naturally occurring neurotransmitter-like substances that have many of the same effects as opiate drugs.

- Indirect dopamine signaling
- Opioids→ dopamine

Dopamine is important for motivation (wanting)



Gluck et al., *Learning and Memory*, 4e, © 2020 Worth Publishers

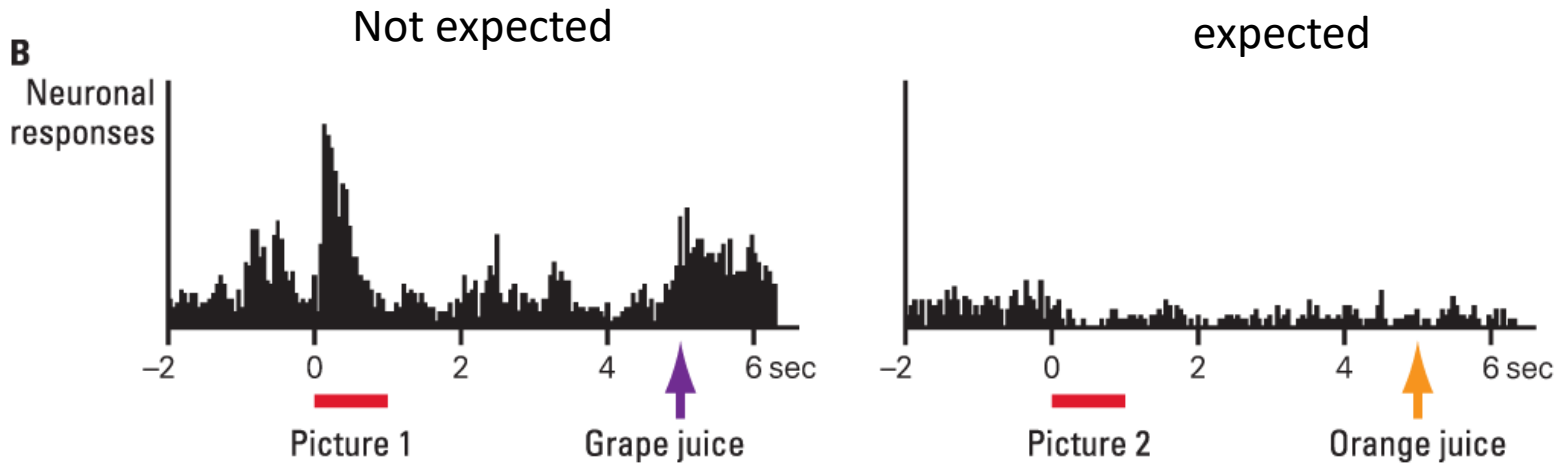
Most rats prefer sugar pellets to rat food. If the food is freely available but sugar pellets have to be "earned" by pressing a lever, control rats (pale blue bars) will spend most of their time working for sugar pellets (A) and will eat relatively little free food (B). In contrast, rats given a dopamine antagonist (dark blue bars) are less willing to work for the sugar pellets and instead settle for eating more of the freely available food.

Dopamine Antagonist – dopamine blocker

Predicting Specific Outcomes

- Neuronal recordings provide strong evidence that the prefrontal cortex plays a role in predicting the outcomes of responses

- dopamine plays a key role in monitoring *reward prediction error*
- Reward prediction error information is critical when an organism is learning how to predict (and act to obtain) future reward



Gluck et al., *Learning and Memory*, 4e, © 2020 Worth Publishers

"wanting" or motivation via the dopamine system

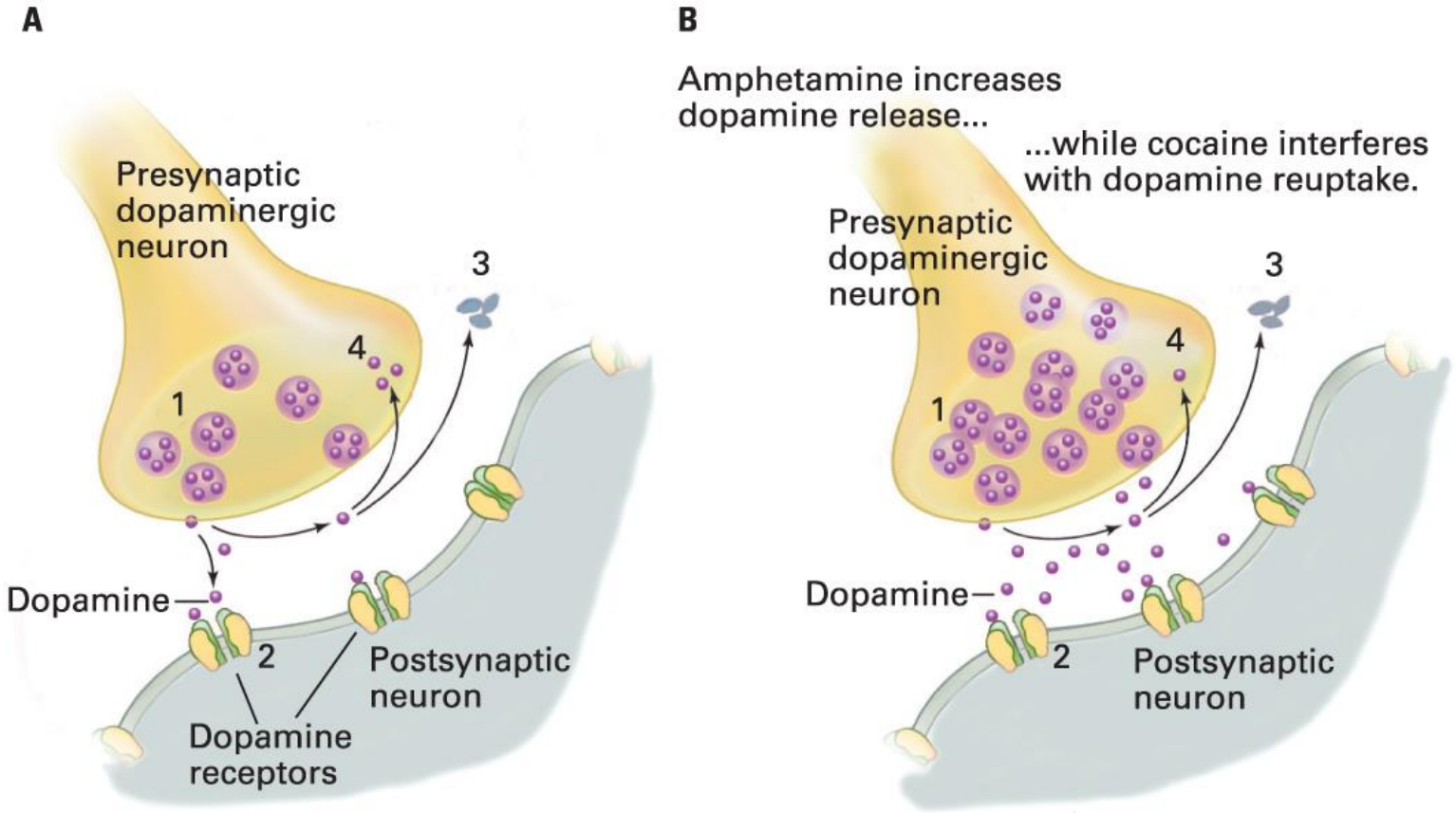
Behavioral Addiction

- **Behavioral addictions** are addictions to behaviors that produce reinforcements or highs, as well as cravings and withdrawal symptoms when the behavior is prevented

- Examples: gambling, sex, social media, gaming

- Behavioral addictions seem to entail dysfunction in the same brain substrates that are affected by drug addictions.

The Effects of Amphetamine and Cocaine on Dopaminergic Neurons

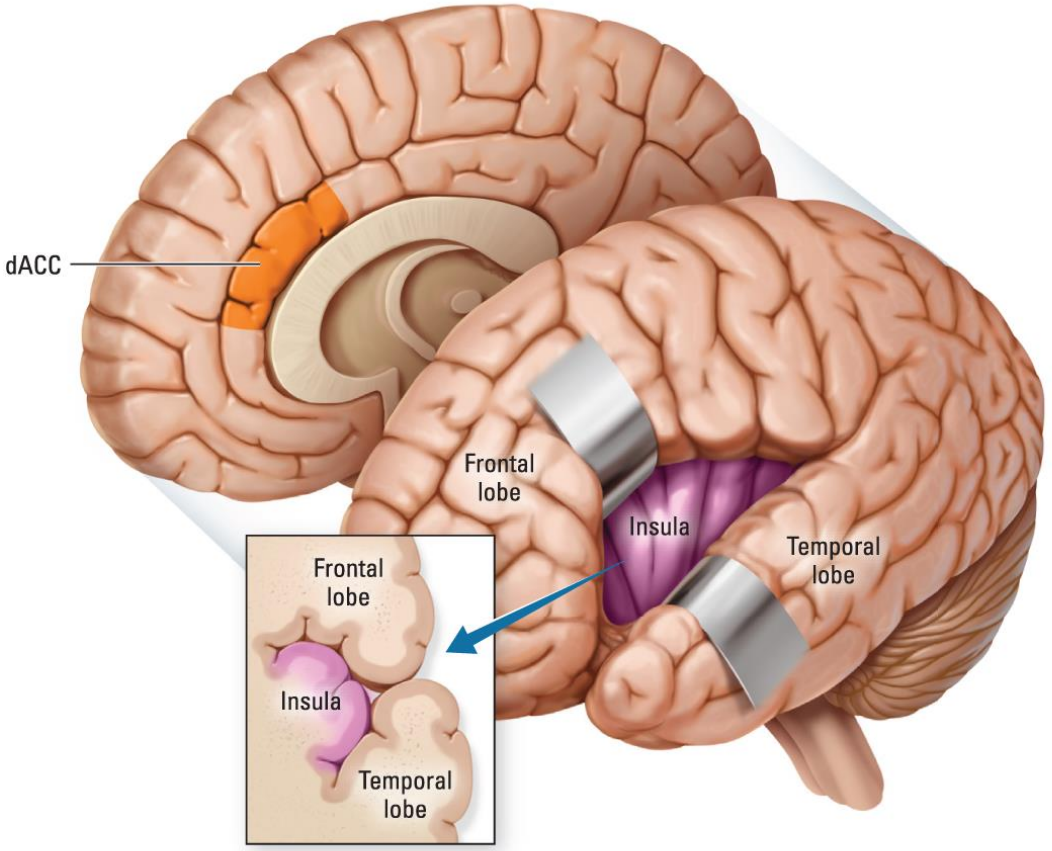


Gluck et al., *Learning and Memory*, 4e, © 2020 Worth Publishers

Insular Cortex (Insula) and Dorsal Anterior Cingulate Cortex (dACC)

Punishment Signaling in the Brain

- **dACC**
- evaluates motivational value of punishing stimuli
- Integrates emotion – reward – action
- links motivational outcomes to behavior
- The dACC shows increased activation when study participants unexpectedly receive a reduced reward



Gluck et al., *Learning and Memory*, 4e, © 2020 Worth Publishers

- The **insular cortex (insula)** is a brain region that is involved in conscious awareness of bodily and emotional states and that may play a role in signaling the aversive value of stimuli

- Evaluates (pain, punishment)
- Relays signals to dACC

OCD and operant conditioning

- Rituals or compulsions
    - Excessive washing of hands
    - Checking repeatedly if doors are locked while leaving the house
    - Always arranging things in a particular order
    - Excessive cleaning of the house

- Rituals help to ease the anxiety and feel better
- Type of conditioning ?
- Solution? therapy?

https://www.sciencedirect.com/science/article/pii/S2949669123000106

https://www.betterhelp.com/advice/personality-disorders/examples-of-operant-conditioning-that-can-help-control-ocd/

Therapy using operant conditioning

- Rehabilitation of stroke patients using VR and operant conditioning

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6712808/