

UNIVERSITÀ DI BOLOGNA
Computer Science and Engineering - DISI

COGNITIONS AND NEUROSCIENCE - NOTES

Student
Alex Periti 0001057793

Academic Year 2021-2022

Contents

1	Introduction	4
1.1	Human brain as a model for AI	4
1.2	Cognition	5
1.2.1	Severed Corpus Callosum Experiment	5
2	Story	6
2.1	Story - the neuroscience side	6
2.2	Story - the cognitive science side	7
2.2.1	Associativism	8
2.2.2	Behaviorism	8
2.2.3	Cognitivism	9
3	Neuroanatomy	10
3.1	Classes of a cell in Nervous System	10
3.2	Regions of a neuron	11
3.3	Synapse	12
3.4	Neuron classification	13
3.5	Neural circuits	14
3.5.1	Principle of connectional specificity	14
3.5.2	Neural circuit: the knee-jerk reflex	15
3.5.3	Parallel Processing	16
3.6	Neural Systems	16
3.6.1	PNS	17
3.6.2	CNS	18
3.7	The Brain	19
3.7.1	Brain Stem	19
3.7.2	Cerebellum	20
3.7.3	Diencephalon	20
3.7.4	Telencephalon	20
3.8	Cerebral Cortex	22
3.8.1	Frontal Lobe	23
3.8.2	Parietal Lobe	23
3.8.3	Occipital Lobe	24
3.8.4	Temporal Lobe	24
3.9	Association cortex	24
4	Neuron Signaling	25
4.1	Signaling unit of the nervous system	25
4.2	Information transfer within a neuron: action potentials	25
4.2.1	Resting membrane potential	26
4.2.2	Action Potential	28
4.3	Information transfer between neurons: synaptic communication	30

4.3.1	Electrical synapses	30
4.3.2	Chemical synapses	30
5	Reinforcement Learning	32
5.1	Non-associative learning	32
5.2	Associative learning	33
5.3	Classical conditioning	33
5.3.1	Classical conditioning in the lab	34
5.3.2	Types of conditioned responses	35
5.3.3	Phases of the conditioning	37
5.4	Neuronal Bases	37
5.4.1	Functional Magnetic Resonance Imaging (fMRI)	37
5.4.2	Lesional Method	38
5.4.3	Neural bases of aversive conditioning acquisition: amygdala	38
5.4.4	Neural bases of aversive conditioning acquisition: hippocampus	38
5.4.5	Generalization	39
5.5	Instrumental Conditioning	40
5.5.1	Shaping	40
6	Mechanism of Reinforcement Learning	42
6.1	Contiguity and Contingency	42
6.1.1	Experimental evidence that compares delay versus stress conditioning	42
6.1.2	Plasticity	43
6.1.3	Challenging the contiguity assumption	44
6.1.4	Learning from unpredicted event	46
6.1.5	Prediction error	47
6.2	Computational Model	47
6.2.1	Rescorla-Wagner	47
6.2.2	Temporal difference	48
6.3	Dopamine	48
6.3.1	Dopaminergic pathways	49
6.3.2	Hypothesis dopamine neuron functioning - Animal	49
6.3.3	Hypothesis dopamine neuron functioning - Human	54
7	Model Free and Model Based Learning	57
7.1	Learning strategies: map vs response strategy	57
7.2	Generations of studies about goal directed and habitual learning	58
7.2.1	Generation 0: cognitive maps vs stimulus-response	58
7.2.2	Generation 1: Goal-Directed vs Habitual actions	61
7.2.3	Generation 2: Actions and Habits in the Human Brain	63
7.2.4	Generation 3: model-based vs model-free computational analyses	64

1 Introduction

1.1 Human brain as a model for AI

Could be a good idea to base the AI on the human brain? For some researchers it may not be a good idea to just base it on the human brain. In the other hand for some of the researchers the human brain is the proof that general intelligence exists. So it may be a good idea to study the human brain, since can inspire the computations and implementations of new models.

Once the neurons start firing they fire and they transmit the message, the message could be binary, yes or no, similar to what also computer does. But of course this is a good model to build computers, but the brain is more complicated than this. Neuroscience can inspire artificial intelligence. Studying the brain can inspire to make new different.

You can see that using the nature, animals and also humans as models can solve technological problems. The reason behind that is that organisms have evolved and now they have well adapted structures. They are fit to survive in their environment through material selection, and so they have learned to solve the problems that are posed by the environment. In the same article [1], not everybody agreed that it's a good idea to model AI on the animal and especially human brain. Because the human brain doesn't implement intelligence in the same way as the computer does. So for example, machine learning practitioner use statistical learning which requires a very large collection of examples, and also a lot of memory capacities. So it's memory and computationally expensive.

Do you need a lots of different trials to learn to do something? The brain instead only need one trial. For example, to learn to do something, we are able to learn in just one trial and also to generalize what we've learned to other contexts and circumstances, because if we were not able it would be really inefficient for survival. We don't know enough about how the brain and the whole secret works, we just know a little bit of a how it works. So we don't know enough to reproduce its structure, in an artificial way.

Living cells, processing incoming information generates not just electrical signal as the computer does, but they also generate chemical signals, actually there are more cells that works based on exchanging chemical signals than electrical ones. Reflecting on this, there is different levels at which you can emulate or mimic the brain. You can closely mimic and reversely engineered the neural circuits of the brain. And that's, for example, what is the aim of the blue brain project, which aims to create a biological detailed representation of the mouse brain. Which is like it's much smaller and less complex, but it's already a challenge.

Create a biological representation means to reconstruct the brain from the micro level to the macro level. From how it works, on a molecular level to the bigger level. So putting neural circuits and neuro structures together, and then you can mimic function. On the other hand you can study the brain but being more interested in the computations that the brain does. And then get inspired by this computation to then transfer them into how you build AI.

What is neuroscience? In general is the study of the nervous system and it's more concerned with the structure, organization and also function on a lower level.

1.2 Cognition

How different senses are integrated, what's multisensory integration?

- Organization: it includes different mental processes and they are all related to the acquisition, storage, manipulation, retrieval and recall of information of different types. It's concerned with the description, explanation and simulation of the thought processes (higher level functioning).
- Attention: allows us to focus on different elements of the environment that are relevant for us.
- Learning: manipulation of concepts, new information and integrate this information with prior knowledge. For learning, memory is crucial because we take information from the past to produce, and to interact with the moment, and produce new knowledge. In order to take the information from the past, we must retrieve, collect and store it.
- Memory: what memory does is: encode, store and retrieve information.
- Action: we use perceive information to interact with the environment.
- Language: the ability to understand and express our thoughts through words written or oral. It allows us to communicate with each other.
- Talk and higher reasoning: takes all the information and processes to engage in a decision making and problem solving. It's the highest level of cognition.

1.2.1 Severed Corpus Callosum Experiment

The left hemisphere is specialized in language, so the structures that enables us to communicate and speak resides mostly in the left hemisphere. What happens is that when left and right are separated the patient can independently move his hands, and draw different shape simultaneously. Furthermore the patient was tested to recognize face from Arcimboldo's pictures, he was able to do it only for left-located images, that were detect by right hemisphere. The patient is not able to understand and express what is shown in right-located images, he can't only through verbally, but he can do it when he's drawing it [2].

2 Story

2.1 Story - the neuroscience side

Tomas Willis (1621-1675)

Was the first anatomist and one of the earliest scientists to realize that brain structure gives rise to brain function. He linked abnormal brain function to abnormal brain structure, brain function and normal brain structure. He was a pioneer of cognitive neuroscience. He discovered that isolated brain damage affects behavior. He coined the term neurology and he discovered lots of different brain structures in nervous system, one that takes his name is the circle of Willis.

Franz Joseph Gall (1758 - 1828)

Birth of **localizationism** which stated that innate faculties were localized in specific regions of the cerebral cortex. Gall hypothesized that if a person used one of the faculties more often than the others, the part of the brain representing that function would grow, causing a bump in the skull. He started to measure people's skulls to infer their personality. He used his study to support, for example, racist claim that people of a certain race had different characteristics (not a scientific approach).

Marie-Jean-Pierre Flourens (1794 - 1867)

Napoleon asked Flourens to prove the validity of the Gall's study about skulls and he showed that indeed certain parts of the brain were responsible for certain functions, however he couldn't find any areas for advanced abilities concluding that the whole brain participated in behavior (**aggregate field theory**)

John Hughlings Jackson (1835 - 1911)

In 1836 Mark Dax reported three patients with speech disturbance, and these three patients had a similar left hemisphere lesion that was found during the autopsy. Jackson had studied patient who had epilepsy. When you have an epilepsy you have electrical activity that's spreading throughout your brain in an abnormal way. He observed there was different parts of the body moving, especially when having seizures and it moved from one part of the body to another in an orderly manner. And this manner would repeat in different patients. Thus he proposed that there was actually a **topographic organization** of the cortex, so that different parts of the body would be represented in different parts of the brain cortex. Jackson's intuition was right, but it's going to take another 100 years before it's actually tested and confirmed.

Paul Broca (1824 - 1880)

Published the results of an autopsy of a patient that was named Tan. His nickname was Tan because the only thing that he was able to say, after having a lesion at left hemisphere, (in particular the inferior frontal lobe) was Tan. Tan had what it's called **expressive aphasia**, he was able to understand language but could not speak (what we know is that a lesion to that area produces those symptoms in most of the patients and this has been replicated)

over and over again)

Korbinian Brodmann (1868 - 1918)

Proved that different regions of the brain are made of different neurons. He divided the brain cortex, in 52 different regions and this field of study is named **cytoarchitectonic**, architectonic of the cell or how cells differ between regions. These different structures also resulting in different functions. Most of the times people refer to these numbers to say which area was interested in the lesion or which area was active during a particular task.

Camillo Golgi (1843 - 1926)

He invented the method to stain a single individual neuron, called **the black reaction**. Camillo impregnated individual neurons with silver chromate, and thanks to this method he started to draw different neurons and also individual neural cells.

Santiago Ramon y Cajal (1852 - 1934)

Santiago gave rise to the neuron doctrine. By using Golgi's technique he realized that the brain is made up of different neurons. Before that it was thought that the brain was made up of only one unitary tissue that shared one extracellular space. Cajal realized that actually there are different neural cells, that make up the brain, and also that the nervous system and neurons are discrete entities, so each neuron is separate from the others. He also realized that neural cells communicate through electricity and the electricity only moves in one direction through the Axon. This discover was so important that lead to the Nobel Prize in Medicine and Physiology in 1906.

At the beginning, scientist believed that the majority of synapses were electrical, but then they realized that majority of synapses are chemical. This is an advantage, wrt the electrical type, because you can use different molecules to do different things and have receptors that respond only to one or another, in this way it's easier to modulate how much of the chemical you can release. Instead the advantages of the electrical synapses is that they are fast.

2.2 Story - the cognitive science side

Cognitive neuroscience exactly tries to understand how brain function depends on integrity and understanding of both structure and function and their relation. It tries to put into relation brain structure and brain function.

Greeks 500 years BC, there were two different views:

- Cardiocentrism: put at the center the cardio, which is the heart. They thought that the heart was the place where the mind was located.
- Encephalocentrism: Hippocrates thought that the intellect and neurological disease originate from the brain.

Rene Descartes (1596 - 1650)

He is famous for the **Cartesian dualism**: mind and body are separable, and the soul controls the body. And then opposite to dualism, there was **monism**: the mind and the body are not separable, and the body produces the mind, so the brain produces our cognition.

In the 17th century there are 2 opposed schools of thought rationalism and empiricism

- Rationalism thought that all the knowledge about the world could be gained only by thinking. So abstract thoughts is the way through which you get to know the truth.
- Empiricism was low level, the truth and knowledge comes from our sensory experience (from our senses). The brain initially it's a blank state. And then through experiences, you shape the brain in order to produce ideas and thoughts. Then when ideas have become associated to each other, you can put them together to form more complex thoughts and ideas.

2.2.1 Associativism

Associationism is the idea that individual experiences are associated with each other and this produces learning.

Hermann Ebbinghaus (1850 - 1909)

discovered that memory processes could be measured and analyzed in a systematic way and with the scientific approach.

Edward L. Thorndike (1874 - 1949) He wrote a monograph called "Animal Intelligence" an experimental study on associative processes in animals. His idea was that experiences are associated together, and they produce learning. He was the first to describe how this happened. Thorndike observed that when a response behavior is followed by a reward, for example food or water, this response would be repeated. In the other hand If the award is not given, then the response is going to go away, it would disappear, and this is called **law of effect**. It's like the conditional reinforcement learning, if you do something and I give you a reward you're going to do it again. If you give something that's unpleasant or painful, you're gonna stop doing what you were doing, and this is an approach to have people stop unwanted actions.

2.2.2 Behaviorism

Behaviorism states that only observable behavior can be studied

John Watson (1878–1958)

He was one of the main researchers in behaviorism. John thought that psychology could be objective only if it was based on observable behavior and mental processes, all talk on mental processes should be avoided. The brain is considered as a blank state, that could be modified through learning and experience.

2.2.3 Cognitivism

Cognitivism state that mental processes can be studied because the brain gives rise to them

Wilder Penfield (1891–1976)

Mental processes can be study because the brain give rise to them, so there is something objective that gives rise to them. Penfield in Canada was one of the most famous scientists of cognitivism, and he invented the **Montreal procedure** to treat epilepsy.

What happens is that when a patient was having epilepsy, he surgically destroyed the neurons in the brain that were the focus of the epilepsy, because they were producing the seizures, they were firing abnormally. So if we destroyed that focus, it would be like saying that you can stop the seizures. Also before destroying the neurons was trying to stimulate different parts of the brain, to see what function they would given rise on. Pendfield was trying to induce electrical currents to specific locations of the brain to see what's the output. By doing so he determined if a part was particularly important for behavior and if it's the case to preserve it.

Observing the result he created different maps of the brain. So Penfield found the evidence for Jackson hypothesis, by stimulating different parts of the motor cortex. He saw that the whole body was represented in motor cortex and the same in the sensory cortex. Hence when stimulating different parts of the sensory cortex, then you would have sensations as if somebody was, for example, touching you.

Donald Hebb (1904–1985)

At the beginning he was studying the effects of brain surgery on the function of the brain with Penfield, and he strengthen this idea that working of the brain can explain behavior and that the psychology and biology of an organism cannot be separated. He stated that the learning has a biological basis. If we associate a stimulus with a reward, your behavior is going to change. Hebb thought that this association had a biological basis, and in particular, that cells that fire together wire together. Thus if two neurons fire at the same time, then their connection, their synapse is going to be strengthened. In other words, neurons can combine together into a single processing unit, and a connection patterns of this unit make up the changing algorithms that determine the brain's response to a stimulus.

Brenda Milner (1918 -)

The memory is located in the temporal lobe, both temporal lobes are important for memory functioning, in particular the hippocampus. Brenda Milner was interested in some of the Panfield's patients which, after neurosurgery, started to complain about memory loss. She provided anatomical and physiological proof that there are multiple memory systems. Brenda also realizes that how much the memory was impaired depended on the extent of the damage. So the more the medial temporal lobe had been removed the more severe was the memory deficit. And again, she also realized that only bilateral resection of the hippocampus resulted in severe amnesia. Instead if you had a truly lateral resection, then you were more fine.

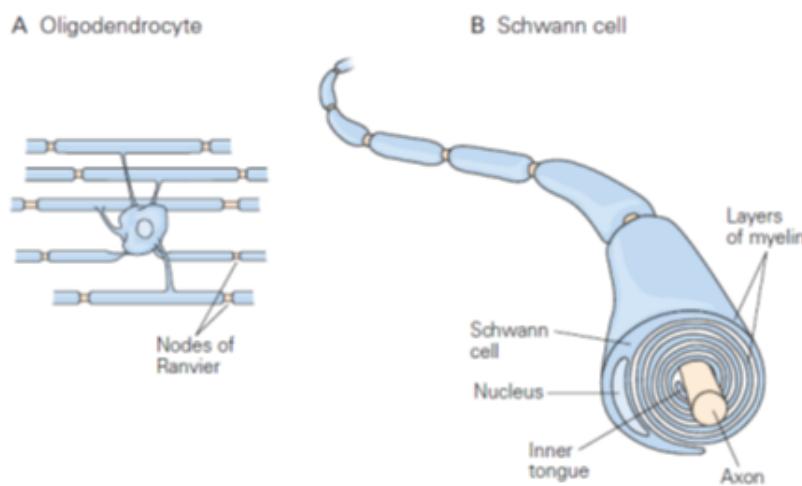
3 Neuroanatomy

3.1 Classes of a cell in Nervous System

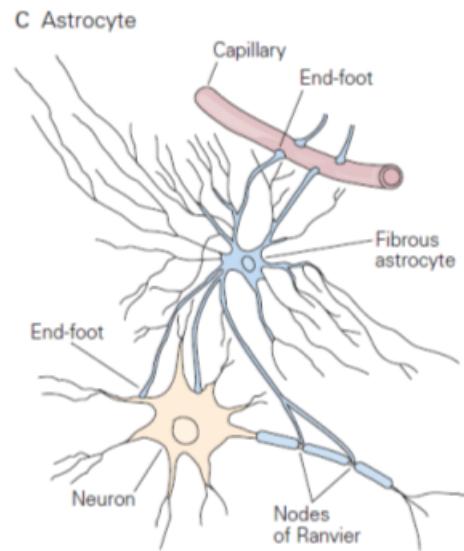
There are actually two classes of cells that form the nervous system, which are the nerve cells, or neurons, and **glial cells** or glia. The neurons transmit signals that you perceive from the environment, **stimulus** to the brain, and then the brain produces an output that travels through neurons, mainly for action. So neurons are the transmitting cells that transmit messages from the periphery to the central nervous system and viceversa. Instead glial cells are supportive cells, that support the neural transmissions and they also are immune system cells, so they help neurons and their nervous system, defending itself from pathogens.

Glial cells outnumber neurons, and there are 2 big classes of glial cells, **microglia** and **macrogliia**.

- Microglia: are immune cells and they are mobilized when there is a pathogen or something that comes from outside the body, and it's dangerous for the nervous system. They become phagocytes during injury, hence they start eating whatever shouldn't be there (like cells);
- Macrogliia: is the most numerous cell type in the human brain (80%) and there is 3 types of macroglia cells oligodendrocytes, Schwann cells, and astrocytes.
 - Oligodendrocyte and Schwann cells: they provide the insulating material around the neuronal axon, by producing myelin sheet that wrapped around the axon of the neuron. This myelin sheet lets signal travel faster and ensure that it doesn't dissipate on the way. Schwann cells are in the peripheral nervous system, while Oligodendrocytes are in the central nervous system;



- Astrocytes: they're called astrocytes due to their star shape. They constitute half of the number of brain cells therefore they can be found in all areas of the brain. They're important for nourishing neurons and in regulating the concentration of ions and neurotransmitters in the extracellular space. Outside of the cell there is fluid that contains various chemicals substances, and astrocytes helps regulate the substances that are in this fluid. Astrocytes can also communicate with each other, although this is not very well understood how signal is transmitted through the neurons;



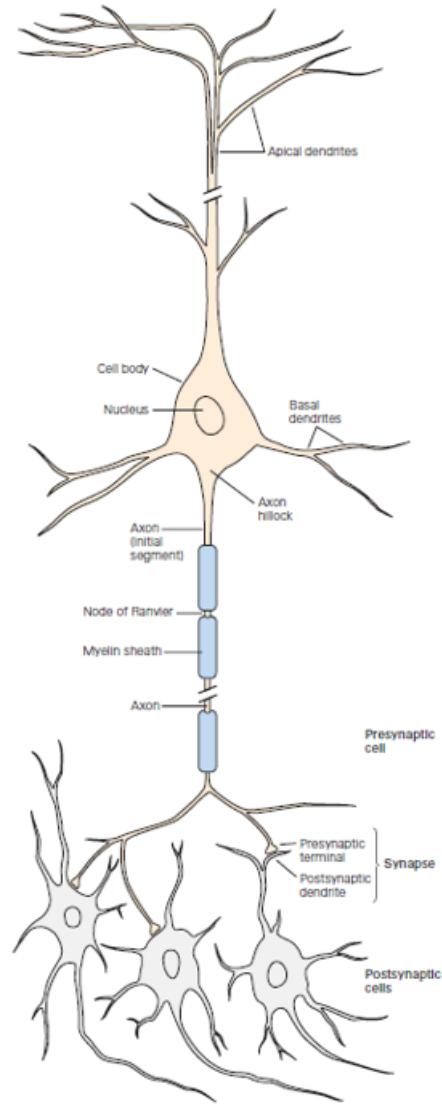
Each neuron receives and also gives rise to 1000 of connections, and some of them are really close to each other, so in some case the accent is really short, instead in other case could be nearly a meter from where the neurons starts.

3.2 Regions of a neuron

In neurons there are four different regions and everyone has one distinct role in generating the signal and communicating with the other cells.

1. Cell body or soma: is the metabolic center of the cell. It contains the nucleus, in which the genetic information reside and the cell's protein are synthesized. Every cells has a nucleus (except for red blood cells). What's important is that the cell body gives rise to 2 processes, **dendrites** and **axon**;
2. Dendrites: receive the input from other nerve cells, and then it arrives to the **axon hillock**, which processes the signal and, if the signal passes a certain threshold, the signal is transmitted;
3. Axon: it can be very short or very long, and it's used to carry the signal through the other neurons. These electrical signals are called **action potentials**;

4. Synapses: is the end part of the axon, is where the message is transmitted through to the other neurons;



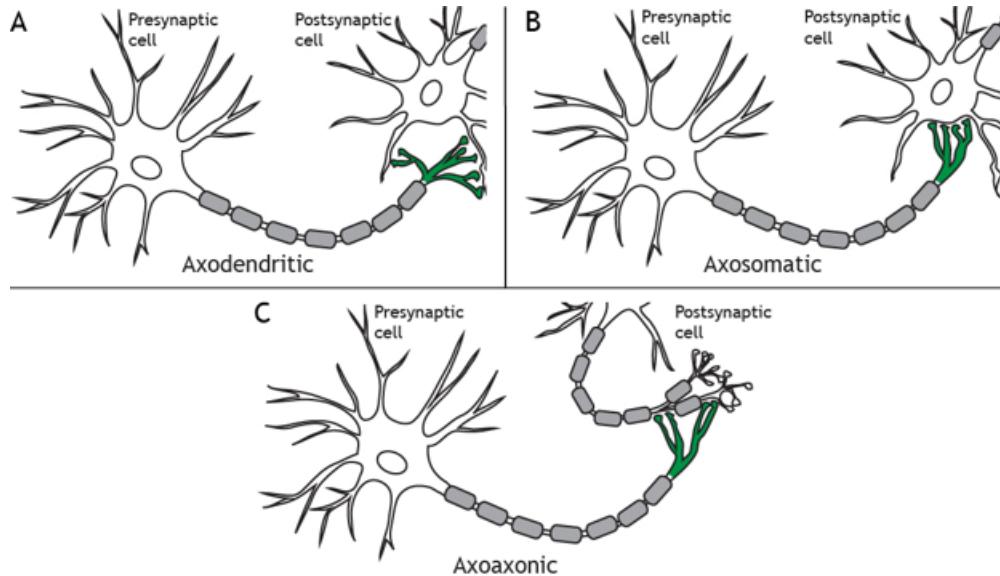
3.3 Synapse

The **presynaptic cell** is the one that transmits the signal to the **postsynaptic cell**. These 2 cells are divided by a narrow space called **synaptic cleft**.

There are three types of synapses:

1. Axosomatic: synapses that are made onto the soma or cell body of a neuron;
2. Axodendritic: synapses that one neuron makes onto the dendrite of another neuron. The most common type;

3. Axoaxonic: synapses made by one neuron onto the synapse of another neuron. Axoaxonic synapses are important to strengthen or reduce or inhibited a signal that is already passed, but just before it's transmitted to the other cell.



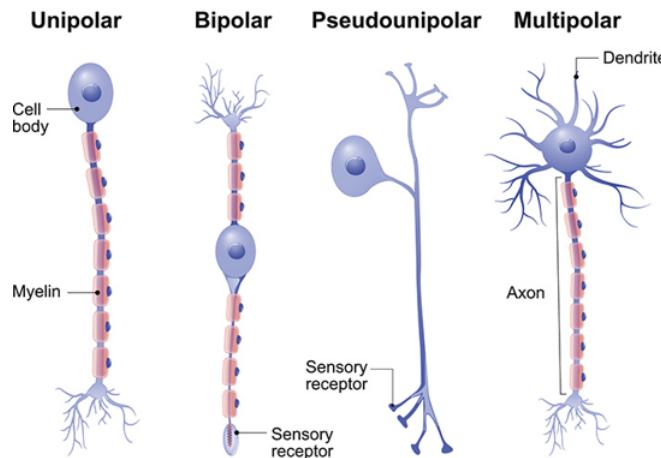
3.4 Neuron classification

Since the structure of the nervous system is related to function, it's possible to classify nerve cells in 3 major class

1. Sensory neurons: receive input both from the environment, and from body. The body information is perceived by the sensory neuron from the peripheral sensors, and from them the info reach the central nervous system. And here the information is processed and generates an output that the majority of the time is a motor response (an action);
2. Motor neurons: they allow the movement by carrying information from the brain or the spinal cord to the muscles and glands. For example, if you are sweating, there is a gland that's sweating and this is because your brain or the motor neuron activates the gland for sweating. This type of information is called **efferent** information, because it goes from the center to the periphery and instead the information carried by sensory neuron is **afferent**;
3. Interneurons: they connect together other neurons. They are subdivided into classes, which are relay and local, depending on how far away they transmit information. Relay have long axons and convey signals over long distances, instead local have short axons.

Neurons can be classified into 3 groups, based on the number of processes that arise from the cell body

1. Unipolar cells: generally doesn't have dendrites. There is only one ramification that arises from the cell body. Unipolar can be a sensor, for example to sense light. They are the most common type in invertebrates;
2. Bipolar cells: it does only 2 processes and these processes are functionally specialized, one is to receive the information and the other to transmit the information;
 - (a) Pseudo-unipolar cells: are a subtype of the bipolar cell. So they start as a bipolar cell but then the two different ramifications get fused into one and emerge from the cell body as a single process that has two functionally distinct segments;
3. Multipolar cells: composed by one Axon and lots of dendrites. They are the most common type of neurons. So usually when you have to draw or represent a new neuron the multipolar cell is what it's used.



3.5 Neural circuits

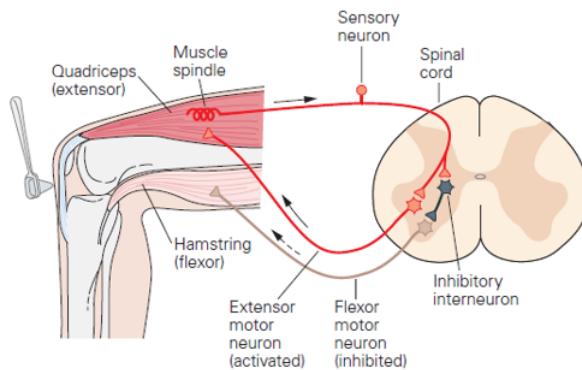
Circuits are groups of neurons, that usually process a specific kind of information. There are different types of circuit processes, for example sensory information (temperature, smell, pressure). Neural circuits are all formed by **afferent inputs**, they take the information in from the peripheral nervous system to the central nervous system and then they evaluate the information that comes in, and then the output produced is called **efferent outward**.

3.5.1 Principle of connectional specificity

Formulated by Cajal says that nerve cells do not connect randomly with one another when they form networks. Instead each cell makes specific connections with other cells with post-synaptic target cells, depending on where the information needs to go.

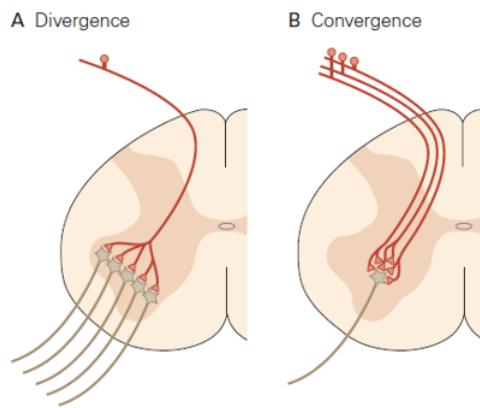
3.5.2 Neural circuit: the knee-jerk reflex

The knee jerk reflex happen when the doctor hit you below the knee, and then your leg it's gonna come up (example of a circuit). When you are hit, what happens is that your quadriceps proceeds information via the muscle spindle, which is a sensor, to the nervous system. It elaborates the stimulus and then simultaneously inhibit the flexor motor neuron of the quadricep and also activate the extensor motor neuron, these 2 muscles are one agonist and one antagonist, this means that when one flexes the other is relaxing, so they have to work in opposite directions to produce movement.



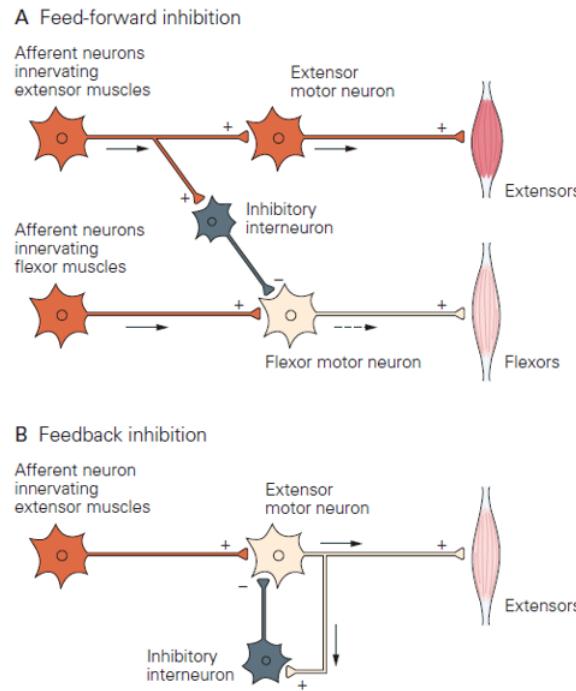
Divergence occurs when a neuron acts on many other neurons. In the figure below the sensory neuron activates lots of other motor neurons.

Convergence is when many neurons converge on one. Converging is useful because you can integrate lots of different formation from different sources into one output. Convergence also ensures that a motor neuron is activated only if a sufficient number of sensory neurons become activated together.



As stated before neurons are both **excitatory** and **inhibitory**. Excitatory neuron is neurons that produces a signal that increases the probability of the postsynaptic neuron to fire, and inhibitory neuron instead reduces the probability that the postsynaptic neuron is gonna fire. In particular there are two types of inhibition feedback and feedforward inhibition:

- Feed-forward inhibition: enhances the effect of the active pathway by suppressing the activity of pathways mediating opposing actions;
- Feedback inhibition: dampens activity within the stimulated pathway and prevent it from exceeding a certain critical level.



3.5.3 Parallel Processing

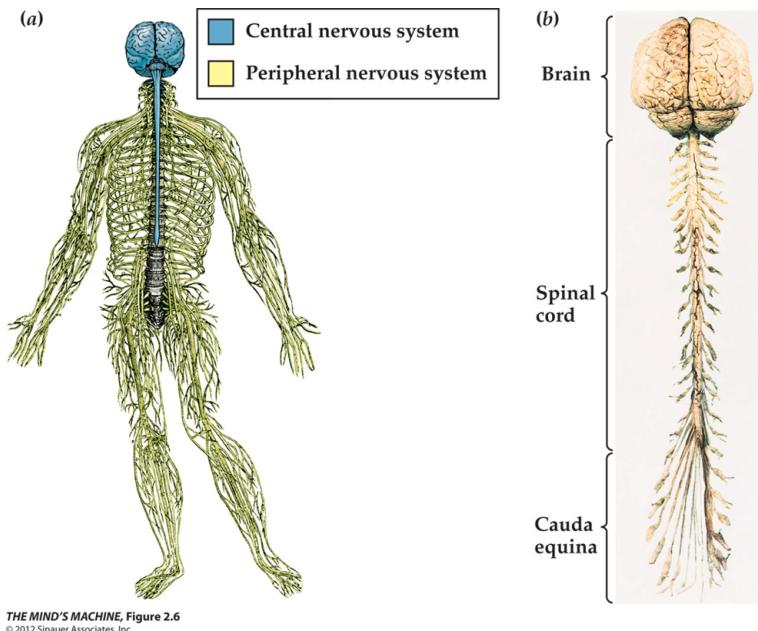
Every complex behaviors doesn't start from a single neuron, but instead from lots of neurons working together. The key components of the neural neural network are the sensory input, then there is an intermediate processing and there is the output, which is mostly motor output. In vertebrates each of these components is mediated by a group of cells or multiple groups only. So multiple group of neurons process the same information and this is called **parallel processing**. Parallel processing is faster and more resilient to changes in the network topology, this increase the reliability of the signal.

3.6 Neural Systems

There are two types of neural systems, central nervous system and peripheral nervous system.

- Central nervous system (CNS): is made up by the **brain** and **spinal cord**, furthermore the neurons that are in the brain and the spinal cord are glial.
- Peripheral nervous system (PNS): is made of the nerves. The cell bodies of cells grouped together are called **ganglia**. PNS delivers sensory information to the central nervous system and carries output information, that are mostly motor commands

from the central nervous system to the muscles. PNS also supplies the CNS with a continuous stream of information that's coming from the outside world, and inside the body. PNS has itself 2 divisions, somatic and autonomic.



3.6.1 PNS

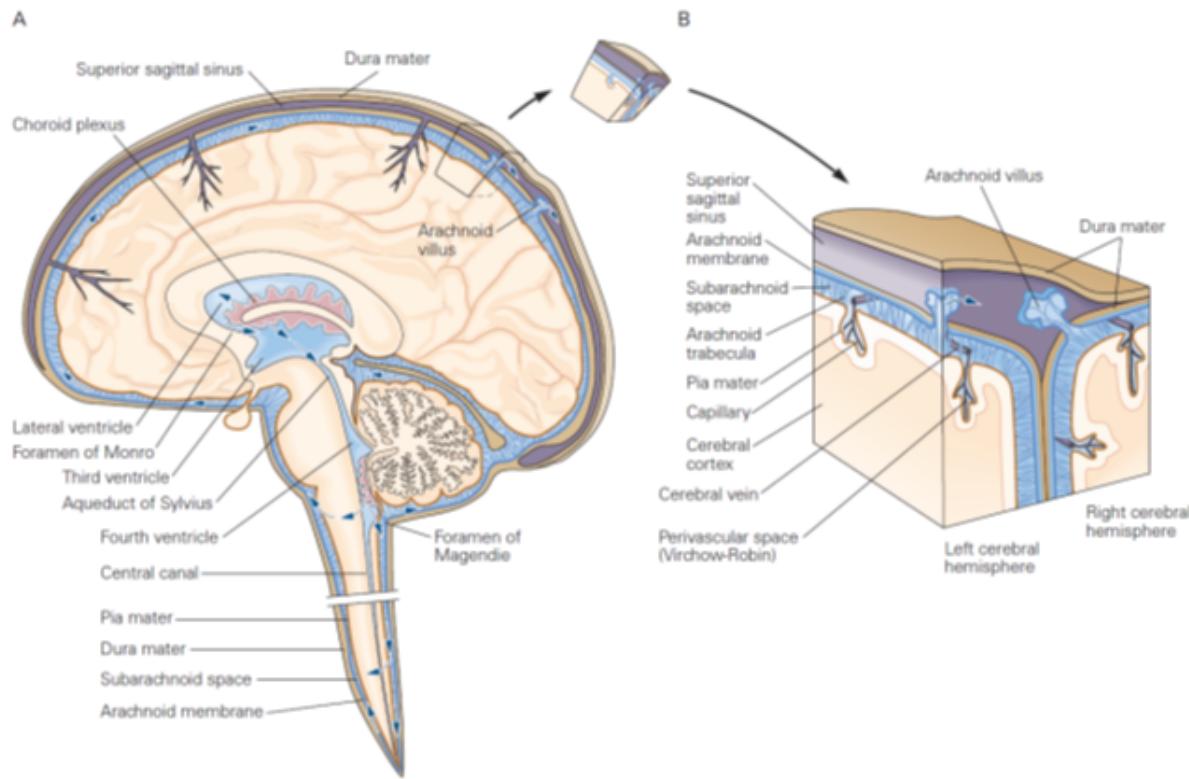
Somatic nervous system is composed by sensory neurons that receive information from the skin, muscles, and joints. The receptors associated mainly provide information about muscle and limb position, as example, where you stand in the environment, in order to move appropriately. What's important is that all these different informations are trasduced into electrical graduated signals in base as the magnitude of the stimulus.

Automatic nervous system controls the visceral sensation, as example the movement of the intestine, the vascular system (blood pressure) and exocrine glands (how much you sweat). It includes 3 systems:

1. Sympathetic system: operates antagonistically wrt Parasympathetic. It prepares the body for action, as example it can dilate pupils, increase the heartbeat to increase the amount of oxygen or even stimulating the adrenal glands to release adrenaline;
2. Parasympathetic system: execute all the action that let the body conserves body resources and restores homeostasis;
3. Enteric system: controls the function of smooth muscle of the gut, so it's important for digestion;

3.6.2 CNS

The CNS is composed by the brain and the spinal cord. Since the brain is really important it's protected by 3 membranes called, **Dura mater** which is the hardest one and most outer one, **Arachnoid mater** and **Pia mater**. The CNS also floats in the cerebrospinal fluid, which is important because it absorbs the shock when you move or hit something, but also it release the pressure generated by the weight of the brain.



The **spinal cord** starts from the base of the brain and ends at the level of the first or second lumbar vertebra. It receives all the sensory informations coming from the environment and send it to the brain, and conduct all the motor signals from brain to muscles. The spinal cord is protected by the vertebral column.

The information, through the nerves, cross from one side of the spinal cord to the other. So information comes in from one side and then across to the other side. It means that one side of the brain controls the other side of the body. Most pathways of the CNS are symmetrical, the **Pineal gland**, it's the only structure that's not bilateral in the brain. So everything else is bilateral and mostly symmetrical.

In the central nervous system neurons are grouped together in different ways

- Nucleus: similar to ganglia they are compact arrangement of nerve cell bodies with similar functions

- Layer: groups of neurons with similar functions

3.7 The Brain

The brain is composed by 6 subdivisions:

1. Medulla
2. Pons
3. Midbrain
4. Cerebellum
5. Diencephalon
6. Cerebral hemispheres or Telencephalon

The brain is mostly symmetrically, everything is duplicated, because if one part is gone, then the other one is still usable.

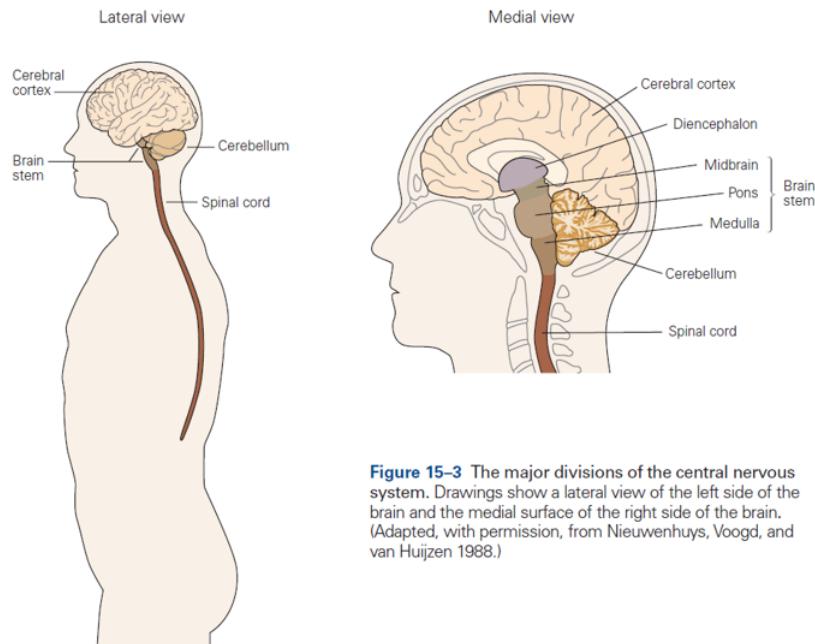


Figure 15–3 The major divisions of the central nervous system. Drawings show a lateral view of the left side of the brain and the medial surface of the right side of the brain. (Adapted, with permission, from Nieuwenhuys, Voogd, and van Huijzen 1988.)

3.7.1 Brain Stem

The brain stem it's made over three other structures, but what's important is that if you damage this structure you could end up in vegetative state. Because the brain stem it's evolutionary older, so it does the most basic tasks, like: respiration, regulation of the sleep, and also it's important for having a basic level of conscious (consciousness). The 3 structures that compose the brain stem are:

1. Medulla: regulates blood pressure and respiration;
2. Pons: connects the brain with the cerebellum. It's composed by 2 parts, ventral portions and dorsal portion, the first it's mainly important for movement, while the second it's involved with respiration, taste and sleep;
3. Mid brain: it's important for linking together components of motor system.

3.7.2 Cerebellum

The cerebellum is important to coordinate the action, for balance and mainly for the procedural memory, that include all long-term memory related to actions that you do automatically.

It's composed by several lobes and they include lot of neurons, indeed cerebellum contains more neurons than all the other subdivision of the brain.

Cerebellum receives the somatic sensation from the spinal cord, as example, balance, posture, so it's important for motion.

3.7.3 Diencephalon

Diencephalon is composed by:

- Thalamus: is essential to link sensory information from the periphery to sensory regions of the cerebral hemispheres. It can be further subdivided into different regions depending on the neurons and the information that arrives there. They're grouped together based on different types of sensory neurons;
- Hypothalamus: it's the link between nervous system and the endocrine system. It regulates the body temperature, hunger, and the circadian rhythm but also the response to threats;

3.7.4 Telencephalon

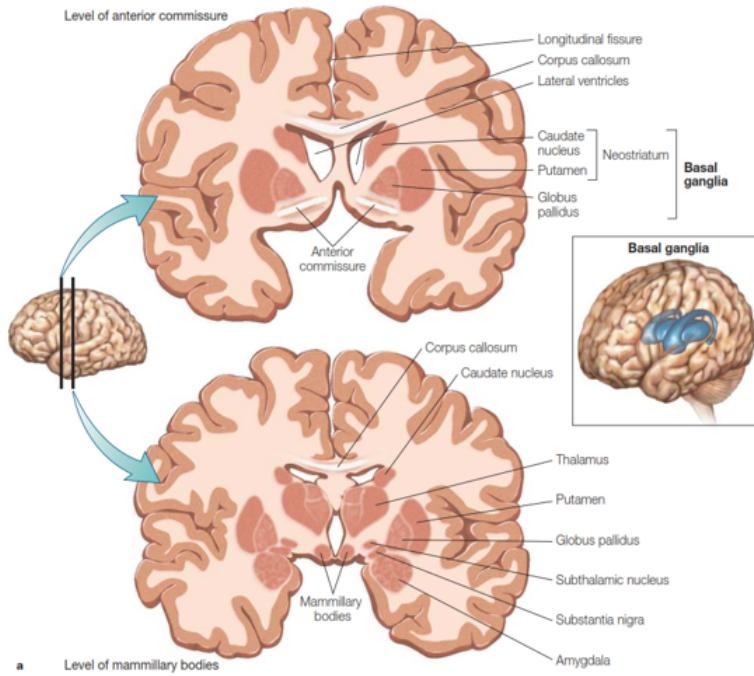
Telencephalon or cerebral hemispheres is the largest part of the human brain, is composed by

- The cerebral cortex
- The underlying white matter
- 3 deep-lying structures:
 - The basal ganglia
 - Amygdala
 - Hippocampal formation

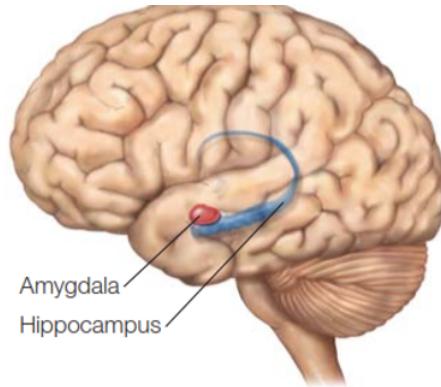
The **Basal Ganglia** is a collection of subcortical nuclei bilaterally located in the brain. They receive the input from sensor and motor areas and send it through the thalamus to the frontal lobe. Basal ganglia is important for making decisions, so when you have different options and you have to decide what option basal ganglia plays a role in this.

Patients with **Parkinson** have a slowing in movement, in particular they have difficulties initiating the movement. So when they have to produce the stimulus by themselves and, for example, they have to queue the movement with another stimulus from the outside, like a growth signal, they have particular difficulties. There is another disease that's also affects the basal ganglia, the **Huntington** disease, which causes a loss of motor function, but it's in the opposite way of Parkinson, patients with Huntington make spontaneous movement that they cannot control. So how is it that the same structure being damaged can lead to two opposite outcomes? In one case a decrease in movement and in the other case an increase in movement? The nuclei of the basal ganglia are connected to each other and include both an inhibitory circuit and an excitatory circuit. So depending on which circuit is affected (by the disease), then you're going to have different symptoms. In the case of Parkinson there is loss of dopaminergic neurons and so these affect the direct pathway and this cause that indirect pathway is over activated, so there is a decrease in movement, while in Huntington is the opposite. In the case of Huntington, it's indirect pathways affected, and so you have over activation of the direct pathway that causes too much movement. The important message is that the basal ganglia are important for movement, in particular for action selection, motor control, and they include both inhibitory and excitatory neurons and depending on which neurons are affected in different diseases, you have different symptoms that can be opposite in different diseases.

Basal ganglia it's important also for **enforcement learning**. When you have a stimulus it conveys some information, and you produce the action or select the action that's better. So that's coherent with the stimulus that you get, and usually you select the action that's going to lead you to a reward or make you avoid the punishment. Basal ganglia do this thanks to **dopamine**, which is a neurotransmitter, which monitors the rewards that comes from the environment, and also the difference between the reward that you expected and the one that you actually got, which is called **prediction error**. Basal Ganglia shapes your actions such that you're going to minimize this error.



Amygdala is a small almond shapes structures that is important for enforcement learning, and it's important for when you detect a stimulus in the environment, because it processes stimulus that can be life threatening or not, and in that case the amygdala it's going to be activated. For example, patients who have lesions through the amygdala had difficulties in perceiving, fearful faces or angry faces, and more generally they are unable to recognize the danger.



The **hippocampus** it's a small structure that is crucial for the memory formation and short-term memory, in particular spatial memory.

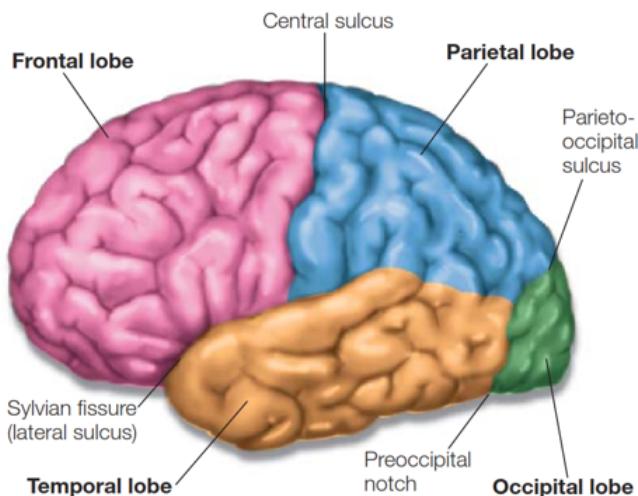
3.8 Cerebral Cortex

The cortex is the outermost tissue of the telencephalon and it's made of sheets of layered neurons. There are two hemispheres, and they are connected to each other via axons that travel

through the corpus callosum. Why do you think we have all these infoldings? If we didn't have all these convolutions, since the total surface area of the brain is 2.4 m^2 it must have convolutions and infoldings in order to make it all fit. Furthermore, axons that make long-distance connections do not follow the structural of the infoldings, and they can travel freely in the white matter, which makes the connections more efficient and more close to each other.

The cortex is divided in the **gray matter** that has the cell bodies and the **white matter** that has the actions. All the information goes through the thalamus. All the information in the cortex, goes from one area of the cortex to the other, and also to other regions of the nervous system.

Cerebral cortex is also divided in different parts, that are recognizable by long different sulcus that divide parts of the brain.



3.8.1 Frontal Lobe

Frontal lobe is divided into:

- Motor cortex: is important for planning and executing movement. In particular, M1 has neurons that directly activates the somatic motor neurons in the spinal cord. And then there is a motor neuron that from the spinal cord goes to the muscle. If I stimulate your M1 electrically, for example, and if I stimulate the one that controls the finger, you're going to see movement in your finger.
- Prefrontal cortex: it is important for long-term movement and long-term planning, this means that it's also important for action selection and for decision making

3.8.2 Parietal Lobe

Parietal lobe receives sensory information from several sources and integrate it together. It also includes the somatosensory cortex, which receives information directly from sensory

neurons about touch, pain, temperature, etc. (what's interesting here is that the central sulcus divides the motor cortex in the frontal lobe from the sensory cortex that it's in the parietal lobe. They are next to each other, but they are divided by a sulcus that's very deep, so they seem close together in three dimensional space. But if you were to put the brain flat, there would be actually very far away)

3.8.3 Occipital Lobe

Occipital lobe play a role in visual pattern matching. The information from your eyes pass through the **optic chiasm** and it's received by the **visual cortex**. Like in sensory/motor cortex there are somatotopic map, the same happens in the visual cortex, and it's called **retinotopic map**.

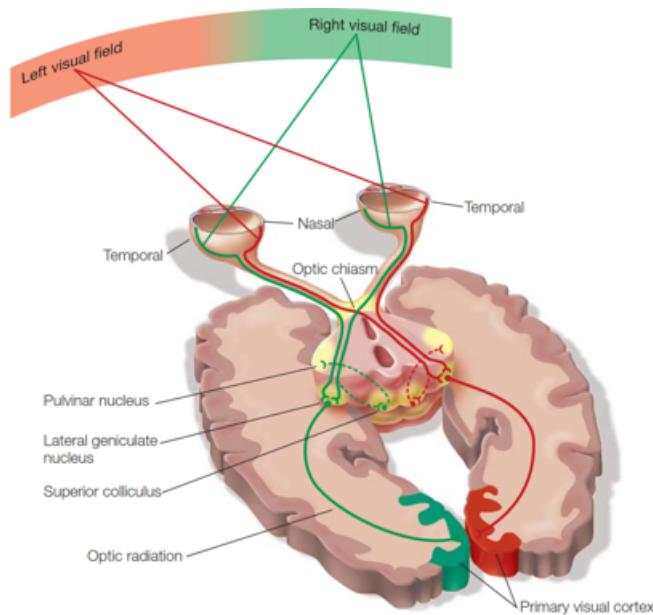


FIGURE 5.23 The primary projection pathways of the visual system.

3.8.4 Temporal Lobe

Temporal lobe includes the **auditory cortex** that processes all different information about sounds, for example pitch frequency. Again, also the auditory cortex is organized in a topographic manner, called **tonotopic manner**, this means that some neurons respond preferentially to some sound frequencies and other neurons to other sound frequencies

3.9 Association cortex

Portion of the neocortex that is neither sensory nor motor, but they are association cortex, so they take all the information that comes from the primary sensory cortices and integrates them so that you can have a whole experience of the outside world, so that you don't process different information separately, but everything comes as a whole.

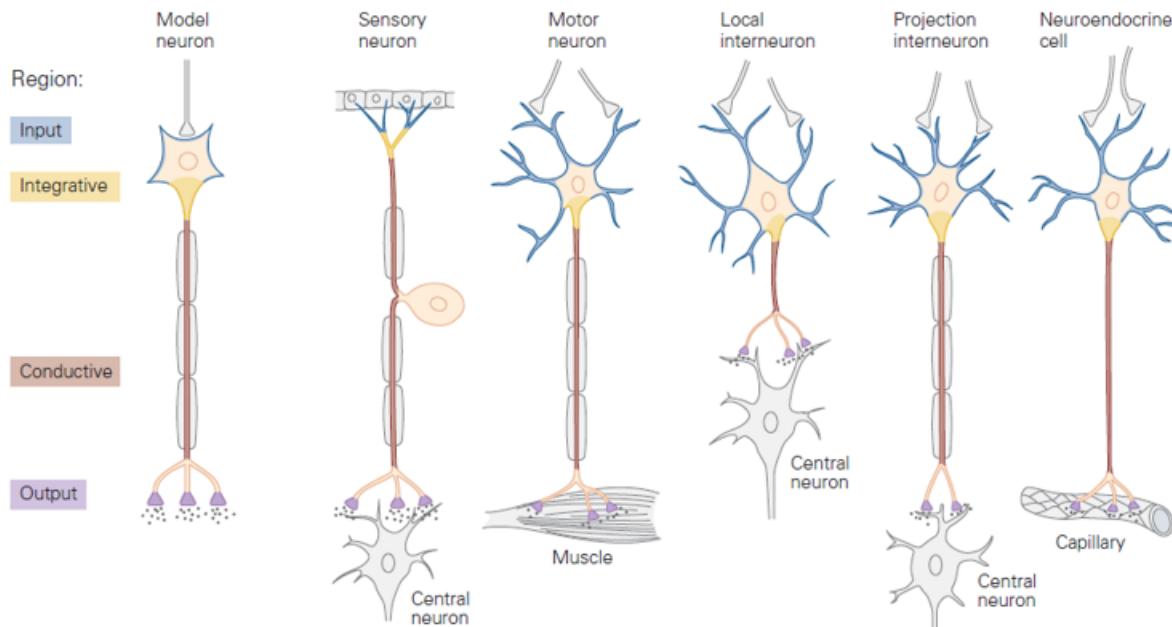
4 Neuron Signaling

4.1 Signaling unit of the nervous system

Information can be transferred within one neuron or between neurons and another type of cell, for example cell muscles or glands. The neuron has an input part, which is the dendrites, then there is an integrated part (the one in yellow in the figure below) which integrates all the inputs that the neuron receives, because the neuron can receive inputs from many different other neurons. So all the input must be integrated and translated into one. The integrative also decides whether the input is strong enough to be transmitted down the axon.

Regardless of the type of neuron, they all have these four different functional parts:

1. Input Signal
2. Trigger signal
3. Conducting signal
4. Output signal



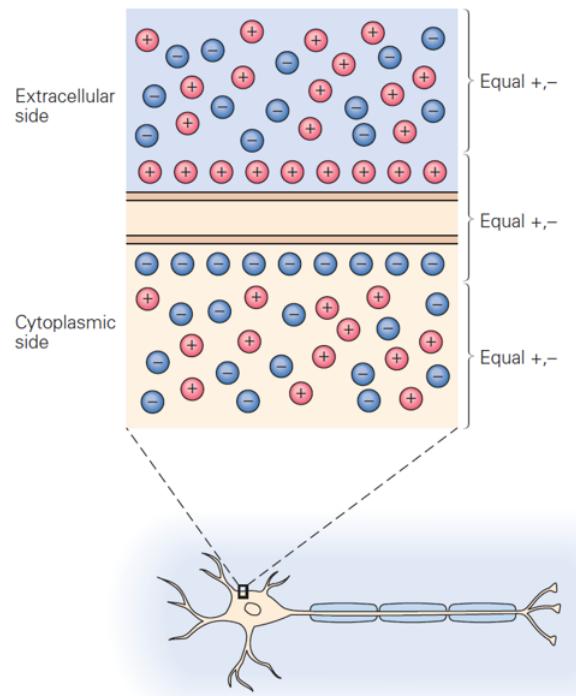
Within a neuron the transfer of information involves changes in electrical states. In a neuron the current flows along the axon creating different neuronal spikes. Between the neurons instead, the messages are transferred at synapses, which can be chemical or electrical.

4.2 Information transfer within a neuron: action potentials

The changes in electrical state of the neuron are possible by producing temporary changes in electrical current between the inside and the outside of the cell.

4.2.1 Resting membrane potential

The cells don't float around by itself, but it's embedded in a fluid, called **cytoplasm**, that's outside the cell. First, we start with a baseline state, called the **resting membrane potential**, that arise from the symmetric distribution of ions across the neuron membrane. It means that, for example, if I stick an electrode inside the cell and I have an electrode outside the cell, there is a difference in electrical potential between the inside and the outside of the cell, and I can exploit this difference in order to have energy float around and make work, transfer this energy into something.

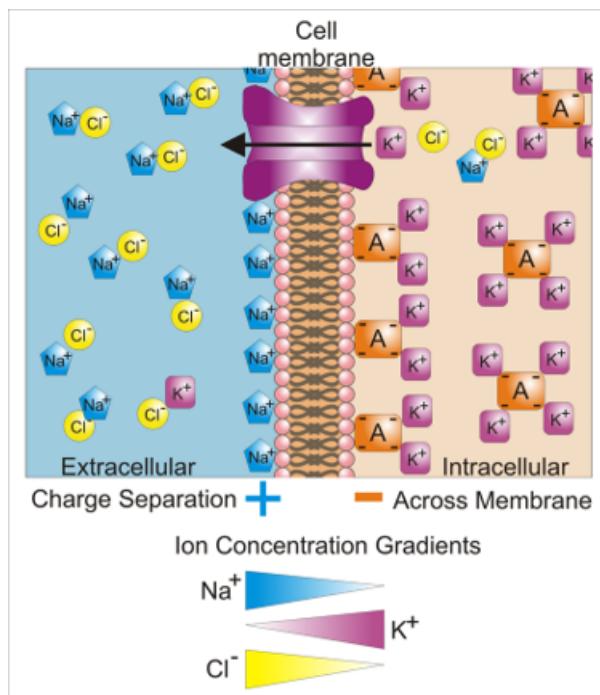


In this baseline state I have different charges, that are just chemicals, that float around in the extracellular space and in the intercellular space, and the main chemicals are sodium and potassium, which are both positive, there is electrical equilibrium in the extracellular and intracellular space. Along the membrane on the extracellular side, there is a lot of positive charges that's sit on the extracellular side of the membrane, and in the intracellular side, there are lots of negative charges that sit around, these two are also in equilibrium.

There are two different mechanism that creates the difference in potential between the outside and the inside, one is called the **sodium potassium pump** and the other are **ion channels**.

The sodium potassium pump works by pumping sodium outside of the cell, in the extracellular space and some potassium inside of the cell. So the concentration of sodium is higher outside than inside of the cell, and the opposite is for the potassium. These different concentrations create a concentration gradient. Then everything wants to go back to equi-

librium, not only in charges, but also in concentration. What it means is the sodium, wants to go into the cell to have the same amount of sodium between inside and outside of the cell and the potassium wants to go out of the cell. But only the potassium can go out of the cell, because there are ion channels, that collect the potassium outside. What happens is that the potassium can go outside of the cell, but since the potassium is positive, if it goes outside, it leaves some negative charges(A) molecules that are not balanced out, and these negative charges molecules position themselves just next to the membrane. These negative charges are balanced by the sodium that is outside of the cell and wants to go in, they also sit around membrane.

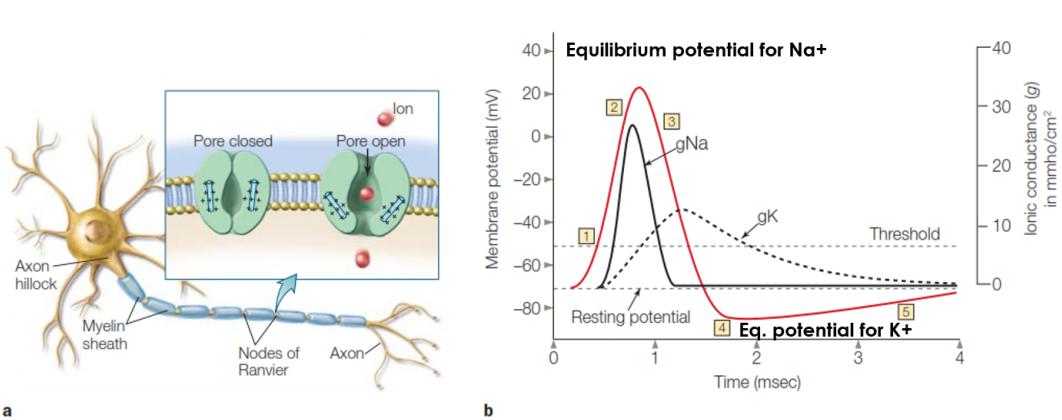


So this is the different potential between in and out, in particular is -70 mV , and represent an energy source that can enable signaling. Thanks to this different potential neurons are excitable, and this means also that the membrane potential can be quickly and significantly altered. And the changes can be both decrease in potential or increase in potential. If the membrane potential decrease, it's called **depolarization**, and means that the negativity decrease and increases the ability to generate the action potential, that is what transfers the message. Otherwise an increase in the membrane potential, it's called **hyperpolarization**, this reduces the probability that the neuron is gonna fire. This changing in the membrane potential it's also useful to distinguish if the outcome is excitatory or inhibitory.

Decrease in negativity \rightarrow the membrane became more positive

Increase in negativity \rightarrow the membrane became more negative

The red line below shows how an action potential looks like.



4.2.2 Action Potential

Starting from the baseline state the membrane potential needs to reach, at least, -55 mv, in order to trigger the **action potential**. Once the threshold is reached, an action potential is generated and what's important is that this action potential is said to be all or none, which means that it has the same shape and the same magnitude in all cells. Furthermore when it's reached the threshold of -55 mv, sodium channel are gonna open. These sodium channels enabled the sodium to go into the cell, and since they are positive, the membrane potential become less negative --> depolarization.

When sodium comes into the cell causes a sudden change in the membrane potential, and this trigger a cycle. The sodium that comes in causes more depolarization and more sodium to come in, and this goes on until another point of equilibrium is reached, around +40 mv, which is the **equilibrium potential for sodium**.

At this point another ion channels open, and these are the potassium channels, this means that the potassium wants to go outside of the cell because of the concentration gradient. When potassium exit the cell the membrane potential starts to decrease until it reaches the **equilibrium potential for potassium**. The potassium channels start to close and the membrane can return to its resting state. The equilibrium potential for potassium is more negative than the resting membrane potential.

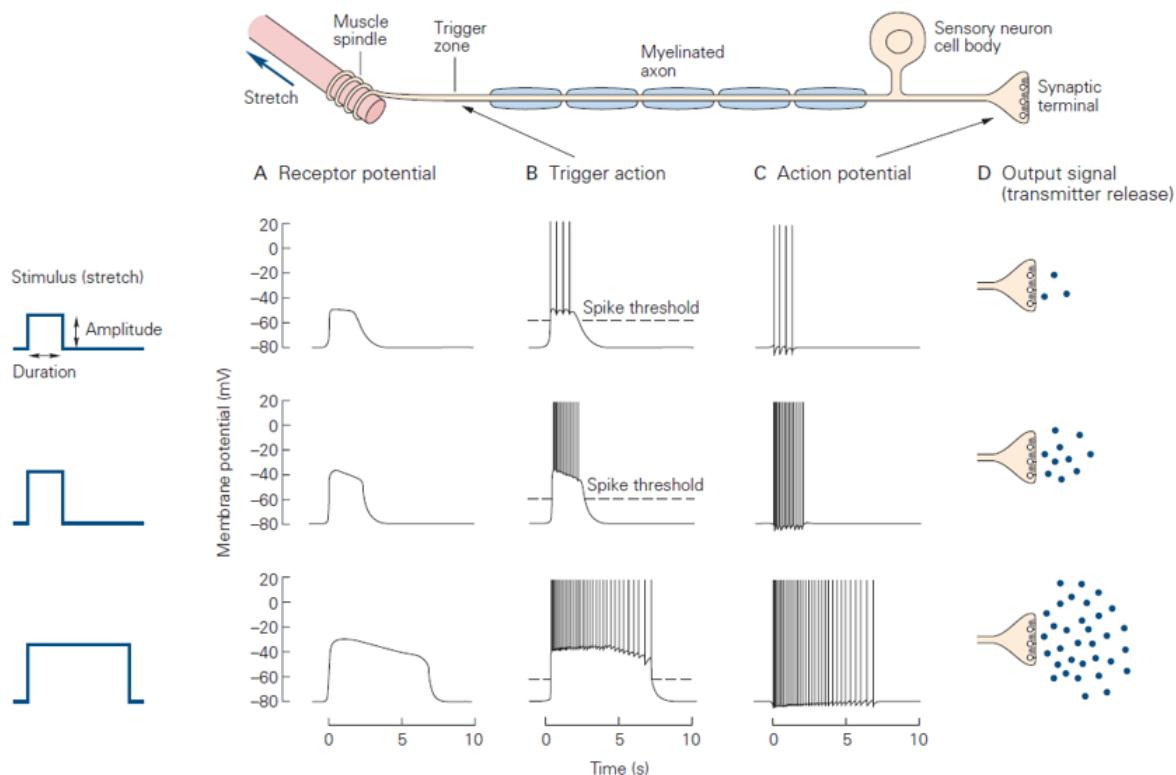
Hyperpolarization is important because it enables the fact that the signal can travel all in one direction, and it also limits the number of action potentials that can be triggered, so the neurons don't start over fire and create for example, seizures.

During hyperpolarization phase the sodium channel can't be open, it means that there is a short time-window in which the neuron cannot fire another action potential, this time-window it's called **refractory period**. All the channels that are near the one that just open cannot open anymore, only the ones that are down through the axon canal can open, this means that the message can travel only in one direction.

The action potential has to travel far without dissipating, this is achieved because the action potential is self regenerative. Each time the channel opens, the action potential is generated again. Action potential also needs to travel fast, it can do it by means of myelin sheets and, in particular, by oligodendrocyte and Schwann cell. The insulating sheets enable the signal to jump from one gap, between the myelin sheet, to another, the gap is called **node of Ranvier**. This behaviour it's called **saltatory conduction**.

The frequency (firing rate) of the action potential is what distinguish one input from another, and the frequency is proportional to the intensity of the stimulus. The more intense is the incoming stimulus, the higher is the firing rate of the neuron.

So the action potential once it's triggered is always the same amplitude regardless any further increase in amplitude of the input signal. What happens is that it changes the frequency depending on the strength and duration of the input signal, and this change in frequency determines how much neurotransmitter is released by the cell.



To summarize, the properties of action potentials are:

1. Threshold for initiation
2. All-or-none nature
3. Conducted without decrement

4. Refractory period

When happen that there is an incoming stimulus, the information is translated into electrical information and then at the synopsis translated into chemical information.

4.3 Information transfer between neurons: synaptic communication

There are two main types of synapses, chemical and electrical, and the 2 work differently.

4.3.1 Electrical synapses

In the **electrical synapses**, the structure where the two cell members touched is called **Gap junction**.

Gap junction has pores which connect one cell to the other, the consequences is that the cytoplasm of the postsynaptic neuron and the cytoplasm of the presynaptic neuron is almost continuous, and also that the two neurons are **isopotential**. There is no separation between the two cells and between their intracellular space, this makes the signal travel fast, and also the input signal cannot change, it remains the same.

Advantages:

- fast transmission
- synchronous operation of groups of neurons

Disadvantages:

- Less plastic than chemical synapses
- Cannot modulated signal
- Less specific

4.3.2 Chemical synapses

In **chemical synapses**, there is no direct structural continuity between the two cells, but the two cells are separated by the synaptic cleft. This space is useful for modulating the chemicals output (neurotransmitter), so the output may be different from the input.

The transmission in chemical synapses depends on the diffusion of **neurotransmitters** across the synaptic cleft, which is the space between the presynaptic and postsynaptic.

Neurotransmitters are chemical substances that bind on the receptor in the postsynaptic membrane of the target cell, so they are chemicals that are found in synapsis and when they are released by the vesicles, they bind on receptors that are on the postsynaptic cell.

Inactivation of neurotransmitters can be accomplished by:

- Active reuptake of the substance back into the presynaptic terminal;
- Enzymatic breakdown or degradation of the transmitter in the synaptic cleft;
- Diffusion of the neurotransmitter away from the site of action;

The effect of the neurotransmitter is determined by not much by the neurotransmitter itself, but mostly by the type of receptors that it meets on the postsynaptic neuron. For example, dopamine can bind either two different types of dopamine receptors. Some of these receptors are inhibitory and some are excitatory, it means that dopamine can have different consequences, excitatory or inhibitory.

5 Reinforcement Learning

Reinforcement learning it's used to describe learning both in what's called **classical** and **instrumental** conditioning.

Reinforcer is a stimulus that causes a change in response, and these responses can be of different nature. Primary reinforcers are those that are biologically programmed to be reinforcing and elicit an unconditioned response. As example, food, pain and sex. And then there are secondary reinforcers that are those that have been learned, and they are frequently associated with primary reinforcers. As example, if I give you money, you're more likely to behave in a certain way. It works also with reward such as social approval or disapproval. After learning that if I have money, I can buy food or I learned that if I have social approval and praise, you're going to be more likely to find a partner, for example, and then have sex.

The responses to the stimuli can be physiological, behavioral or a change in subjective experience. Usually we are mostly interested in physiological changes because it represents how your body reacts to the stimuli, and also in behavioral changes, for example, changes in a reflexive behavior (motor reflexes). The stimulus causes a change in response and this change can be in both directions, it can be in an increase in response or a decrease in the probability of a meeting response.

Learning is an enduring change in response or behavior that is a consequence of experience. There are broad classifications of learning:

- Non-associative
- Associative

5.1 Non-associative learning

Basic type of learning which is also known as non-associative learning. In this case there is only one stimulus that is presented to the subject, it has to learn something about only one stimulus. This can happen also in one trial, or sometimes you need the subject needs to be repeatedly exposed to the stimulus. There are two types of non-associative learning one is **habituation** and one is **sensitization**, and the two have opposite consequences.

- Habituation: In habituation you decrease an innate response to a stimulus that is presented repeatedly. So I present the stimulus lots of times, at the beginning you have a high response, and then your response decreases, when I repeat this stimulus.

Why habituation can be adaptive? The adaptive nature of habituation is that the subject stops responding to stimuli that are not dangerous for him, so he gets used to this stimulus and he doesn't keep responding again and again;

- Sensitization: In sensitization the mechanism is similar to habituation, but in this case there is an increase in an innate response to a stimulus that is presented repeatedly;

5.2 Associative learning

Associative learning is the way in which animals predict events. The animal learns about the relationship between:

- A stimulus and an outcome --> Pavlovian or classical conditioning (experiment of Watson on little Albert);
- A behavior and an outcome --> instrumental or operant conditioning;

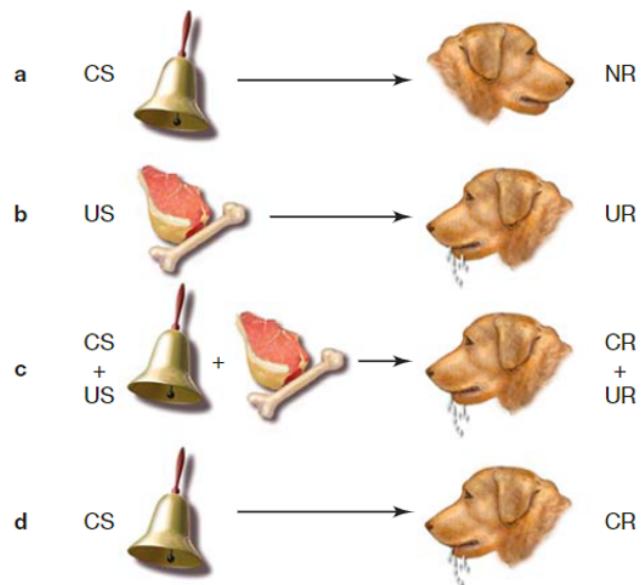
Associative learning is constrained by the biology of the organism, so there are some responses that are more easily to change depending on the way I perform the associative learning. Take as example a stimulus relevant for the survival, as food, and take an aversive stimulation, suppose that you eat something and then after a couple of hours you feel sick. So the next time you see the food or the drink, you don't want to have it anymore, and this could easily happen with only one trial, so it's very strong. What's interesting is that it occurs only when foods or drinks are associated with nausea or sickness. So if the food, for example, is associated with an aversive stimulus, but this stimulus doesn't produce sickness or nausea, then **taste aversion** doesn't occur.

5.3 Classical conditioning

Classical conditioning also called **Pavlovian conditioning**, named after Ivan Pablo, who studied the digestive system of dogs. He realized that when he presented a sound to a dog, at the beginning, the sound didn't have any effect on the salivation of the dog, while the opposite happened when Pablo showed food. Then he paired the sound of the bell with the food, and after lots of trials, the dog starts to salivate when he hears the bell. This happens because when you use the sound, the dog is expecting that the food is gonna come, he anticipates the response. In this scenario the sound is called the **condition stimulus(CS)**. At the beginning there is no response to the stimulus. Then there's an **unconditioned stimulus(US)**, which is a stimulus that instead is gonna get an innate response, called **unconditioned response(UR)**. Non-conditioned response is a response that occurs when I present you an unconditioned stimulus. The stimulus is biologically programmed to elicit a certain response.

When pairing the 2 stimulus, food and sound, I get both **conditioned response (CR)** and UR, and after with only the sound stimulus I get the CR. This is called the **appetitive conditioning** because the US is the reward.

There is also another form that it's called **aversive conditioning**, and that's when the unconditioned stimulus is something noxious.

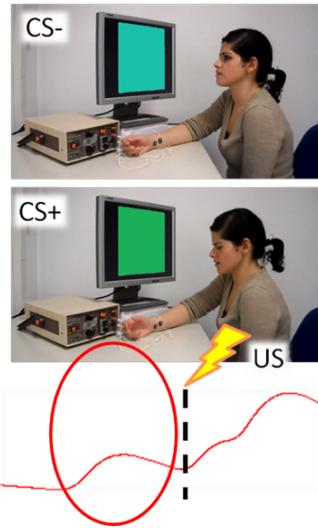


The condition and unconditioned stimulus can be paired different. So for example, I can present the sound of the bell, the condition stimulus, and just before it disappears, I can give you the unconditioned stimulus, for example, give food. Or I can present the sound of the bell. Then the bell disappears and after sometimes I give you the unconditioned stimulus.

5.3.1 Classical conditioning in the lab

Scientists can apply fear conditioning to study the neural mechanisms of learning and memory. Usually fear conditioning it's combine with functional magnetic resonance imaging to investigate the brain regions involved in acquiring and maintaining the conditioned fear response. Scientist study fear conditioning for treatment of phobias, anxiety disorders or also post-traumatic stress disorders. The scientists think that fear conditioning and associative learning is one of the mechanisms through which these types of disorders develop. So by studying these in healthy participants and also in participants with the disorder, they're trying to understand how to help people affected by this disorder.

Acquisition



The experiment has the purpose to see if there is a response before the shock is given. At the beginning the subject only see a response when I present the shock, which is the unconditioned response. But when the green square was paired with shock, as soon as the subject see the green square, her skin conductance goes up, and she started to sweat.

5.3.2 Types of conditioned responses

There are different types of conditional responses:

- Physiological: skin conductance response is a type of physiological response;
- Behavioral: running away if someone shot at you;
- Change in subjective experience: like the experiment of McDonald icon done by the prof;

For example sweating in particular is controlled mostly by the sympathetic nervous system. But we can measure also other physiological responses, like pupil dilation. We can see if pupil dilation changes for the condition stimulus compared to control stimulus, and the pupil dilation is controlled by both sympathetic and parasympathetic nervous system. Another response that's measured usually is the changes in heart rate, when you're scared your heart beats faster. That's also comes from the interaction between sympathetic and parasympathetic nervous system.

Just to summarize, we can measure different types of behavioral responses or also psychophysiological responses, and ask participants to report how they feel or how much they like the stimulus, because the scientist can learn what parts of the nervous system are responsible for each type of behavior, and they are also interested in having all these different measures, because they tell us different things, but also because humans can lie and thus

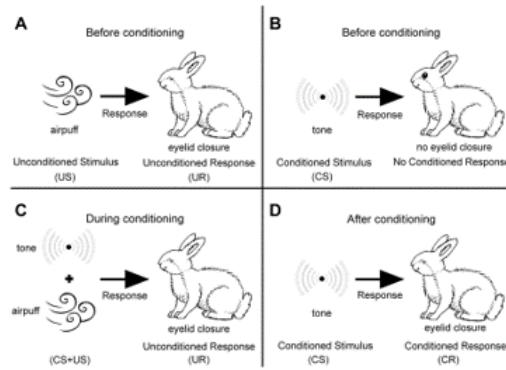
alter the test results.

There are different types of physiological responses. Considering all these responses, once you've learned what happened before, so when you anticipated the stimulus, is like your brain is making a prediction. And translate this prediction into a physiological response.

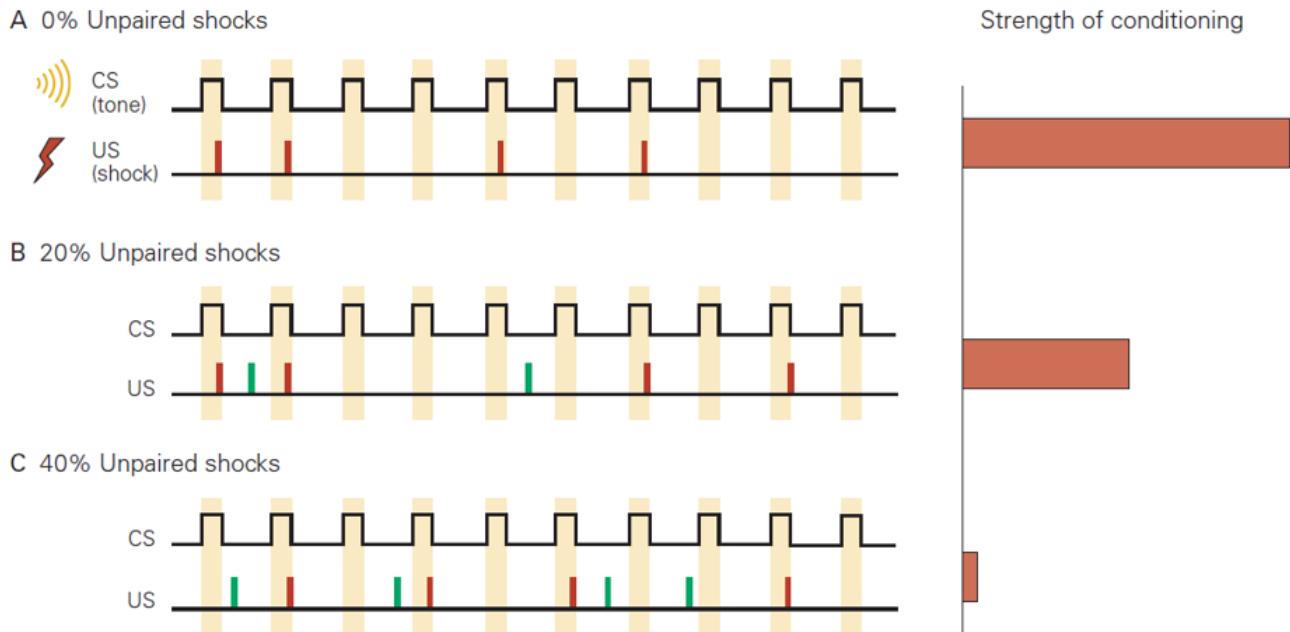
An associative learning is important because I can anticipate the arrival of the reward or punishment, and I can act before these arrives with an adaptive response. So if it's a reward, I can prepare to catch the reward. If it's a punishment, I can prepare myself to run away. So for example, we have a rabbit and an aircraft that's directed to the rabbit, that induces an eyelid closure. And then I also have a ton, if I play the tone nothing happens. Then during conditioning I paired the tone with the aircraft. Hence after the conditioning, if I just play the tone the rabbit is gonna close the eyes.

eyelid closure --> conditional response

tone --> condition stimulus.



The strength of the conditioned response depends on the correlation between the condition stimulus and the unconditioned stimulus. In the example below we can deduce that more random is the association between the tone and the shock and less is the strength conditioning, because would be counterproductive to spend energy to elicit a conditioned response if most of the time actually there is no US after the tone.



5.3.3 Phases of the conditioning

There are 2 main phases of conditioning, one is the **acquisition** phase where I learn the association between the condition and the unconditioned stimulus. So during acquisition what happens is that there is an increase in probability that the condition stimulus is gonna elicit the response.

Then there is **extinction** that happens after I've acquired an association, I stopped pairing the stimuli with the shock, so I still see the stimulus, but there is no more shock. At the beginning I still respond to the stimulus but later, there is no more need to respond to the condition stimulus. Extinction can be adaptive because I shouldn't spend time and energy on producing a response that it's no more useful. I don't forget that once the stimulus was paired with the unconditioned stimulus, but I learned that there is no more need to produce the condition response. Another way to extinguish a condition response it's also the simple passage of time.

5.4 Neuronal Bases

5.4.1 Functional Magnetic Resonance Imaging (fMRI)

Functional MRI is a neuroimaging techniques used to show brain activity, it works by detecting changes in the oxygen levels of the blood in the brain. With this technique you don't measure directly the activity of neurons, but you infer the activity of neurons depending on what happens to the blood around them.

So it's not a direct measure of the neural activity, but you infer how much that part

of the brain is active, assessing the ratio between how much oxygenated and deoxygenated blood there is around.

And the other important thing is that fMRI provides **correlational evidence**, so you can only know that those areas are involved in a certain behavior or response, but then it doesn't tell you if those area are necessary for that.

5.4.2 Lesional Method

There is another technique the **Lesional method**, which study the consequences and behavior after having a lesion to a certain part of the brain. This technique is different from fMRI because it provide the evidence that tells you if an area of the brain is necessary for a certain response. The lesion can be either natural occurring, for example from after a tumor or artificially occurring, as example the surgeon caused the lesion to treat for example epilepsy.

5.4.3 Neural bases of aversive conditioning acquisition: amygdala

Thanks to lots of studies, there are evidences that the amygdala is mostly involved in the acquisition phase, so it's important to produce the conditioned response, while the **ventromedial prefrontal cortex** instead is important to produce the inhibition of conditional response during the extinction.

When there is a conditioned and an unconditioned stimulus, the amygdala receives both directly from the thalamus and indirectly from the cortex. Then the amygdala sends a response to other parts of the brain.

Amygdala presents what is called the **low road** and the **high road**. So the information goes from the thalamus to the amygdala via the low road which is direct, and induces an immediate response, and information from the thalamus also goes through the visual cortex and I become aware of the danger (scenario where i saw i bear). Afterwards the information again goes to the amygdala and this is called the high road. People who have a lesion to the amygdala are not able to produce the conditioned response, so we can conclude that the amygdala is both involved and necessary to produce the conditioned response.

Thanks to a study from 1999 there are evidence that indeed amygdala is important for the acquisition phase, but also that patients with the ventromedial prefrontal cortex lesion have conditioned response similar to the healthy patients so it's is important for extinction and not for acquisition. This theory was disproved by a study conducted by Professor Di Pellegrino (flex della Prof Starita)

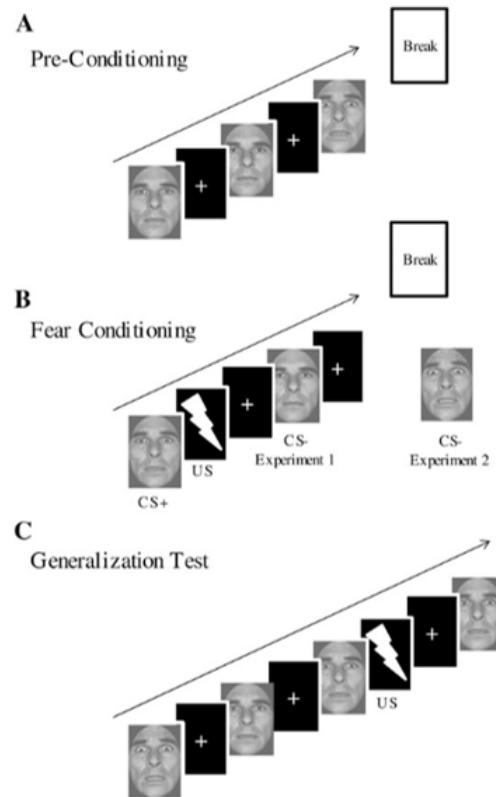
5.4.4 Neural bases of aversive conditioning acquisition: hippocampus

Instead with patients that have hippocampus lesions, they are not able to tell you which stimulus shocked them, but they do produce different skin conductance response. When

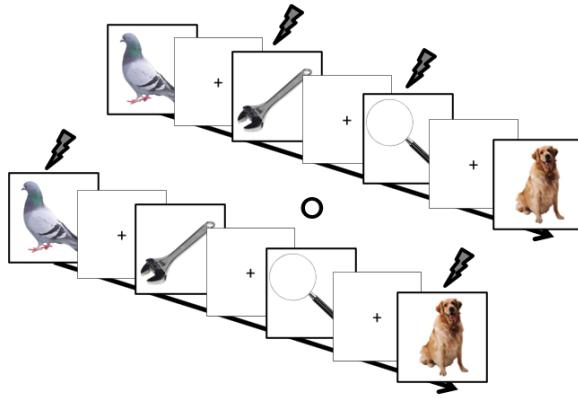
we find that a region is responsible for one process, but not enough to matter and another region is showed the opposite pattern, we call this as **double dissociation**.

5.4.5 Generalization

Generalization happens when I not only respond to the condition stimulus but I start to produce conditioned responses also to stimuli that are similar to the condition stimulus. These generalizations can happen up along different dimensions. One is **perceptual similarity**, so the stimuli are similar under some perceptual properties.



Another type of dimension is the **conceptual similarity**. So here for example, are present conditioned objects, so different types of objects were paired with the shock, while animals were not paired with shocks. The subject understands that every time there is an object or a tool, he should produce a condition stimulus. But every time the subject sees an animal, he shouldn't produce a condition stimulus.



5.5 Instrumental Conditioning

Instrumental conditioning involves associating an action with an outcome.

Thorndike's law of effect “*Of several responses made to the same situation, those which are accompanied or closely followed by satisfaction to the animal will, other things being equal, be more firmly connected with the situation, so that, when it recurs, they will be more likely to recur; those which are accompanied or closely followed by discomfort to the animal will, other things being equal, have their connections with that situation weakened, so that, when it recurs, they will be less likely to occur. The greater the satisfaction or discomfort, the greater the strengthening or weakening of the bond*”.

5.5.1 Shaping

It's very hard that the person or the animal execute complex behavior by itself, but it's possible to reach this goal using the **shaping** technique. If we want to train a very complex behavior in order to have the animal successfully perform it, we break the behavior in smaller steps, and we gradually reinforce each step, which are successive approximations of the behavior.

The **nature of the reinforcement** shapes the behaviour.

Different punishments or rewards have a different effect. The reinforcement increases the probability that I'm gonna perform that behavior again and the punishment instead decreases the probability of behavior performance, so the nature of their enforcement shape behavior.

Positive --> means that we're going to deliver something.

Negative --> means that we're going to take something away.

So if it's a **positive reinforcement**, we're going to deliver a reward. If it's **positive punishment**, we're going to deliver a punishment. If it's **negative reinforcement**, we're going to take out, and if it's **negative punishment**, we're going to take away or omit. For example a positive punishment could be a ticket(multa) and a negative punishment could be the beeping of the car if you don't wear the security belt.

The frequency of the reinforcement also shapes behavior.

Continuous reinforcement, every time you do something, you're gonna get either reward or punishment, so you're gonna be reinforced or punished. This is most effective when you're trying to teach something new, because you don't know what you have to do, but it extinguishes fast.

Partial reinforcement, the behavior is going to be acquired more slowly, but the response is more resistant to extinction, even if I don't give you reward often. There are 4 different types of partial reinforcement:

1. Fixed-ratio, a response is reinforced only after a specified number of responses. This schedule produces a high, steady rate of responding with only a brief pause after the delivery of the reinforcer;
2. Variable-ratio, a response is enforced after an unpredictable number of responses. This also creates a high stated rate of responding and knew the risk is also to become even more compulsive;
3. Fixed-interval, also in this case we have the word fix, so it means that the food is going to come after a predictable interval of time. The animal can also encode the time, so it can encode when the the food is going to come, this means that there is a high amount of responding, especially near the end of the interval;
4. Variable-interval, the reward it's going to come after an unpredictable interval of time, this schedule produces a slow, steady rate of response;

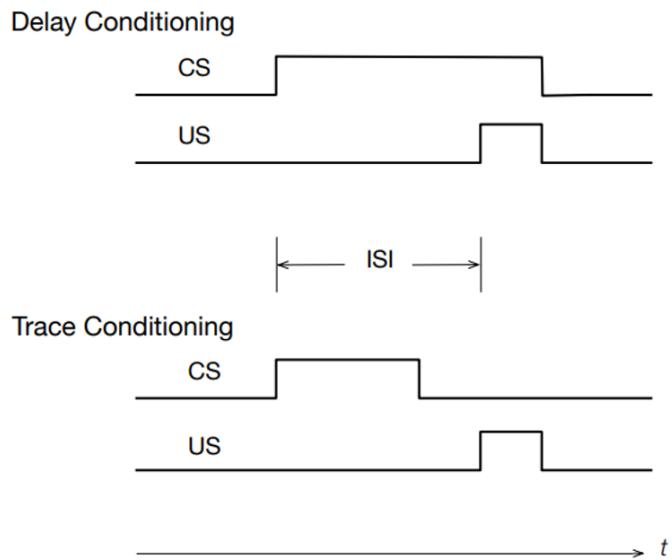
6 Mechanism of Reinforcement Learning

6.1 Contiguity and Contingency

Associative learning is learning about contiguity and contingency:

- Contiguity, means closeness in times between the stimulus (or behavior) and the outcome. Stimuli that are close to one another in time become associated;
- Contingency, causal relationship between stimulus/behavior and outcome. When one stimulus depends on the other, they will become associated;

Initially psychologists and scientist thought that contiguity was the crucial condition for learning. Then they realized that actually there's also contingency that's important. What it's crucial it's not as much how that the two stimuli need to be closing in time, but it's more about the causal relationship between the stimuli and the outcome.



The difference between the 2 stimulus is that in one the US comes when the CS is still active, while in the trace one stimulus comes after the other.

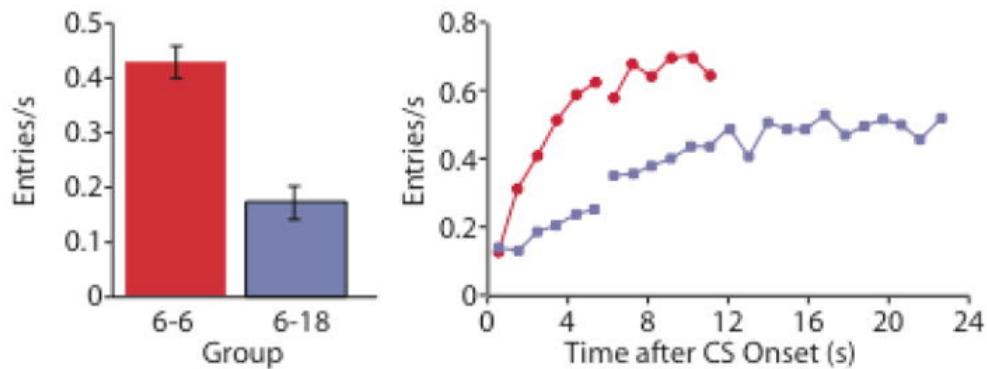
6.1.1 Experimental evidence that compares delay versus stress conditioning

In this scenario scientists experimented the increase of the length of the time between CS disappearance and appearance on 2 groups of rats. Result gives evidence that an increase in this interval weakens the conditioned. We can see that on average the red curve is above the purple curve, so on average there is more entries for the six second group. At zero my CS appears the two groups start at the same time.

Then the red group increases the entries and reaches a plateau. So the maximum increase it's around the time when the CS disappears and then it stays more or less the same. The

same happened also in the purple group.

Basically here we can deduce that if half a second passes between CS and US, there is a strong condition response and increasing condition response. But then the more and more time passes between conditioned and unconditioned stimulus than my condition response decreases, because it becomes harder to associate the two stimuli.



Only closely spaced representation of CS and US allow for the mental representation to be simultaneously active, so the CS representation and US representation need to be active together to form the association.

6.1.2 Plasticity

Plasticity is about how learning and experience can modify connections between our neurons and the way they change.

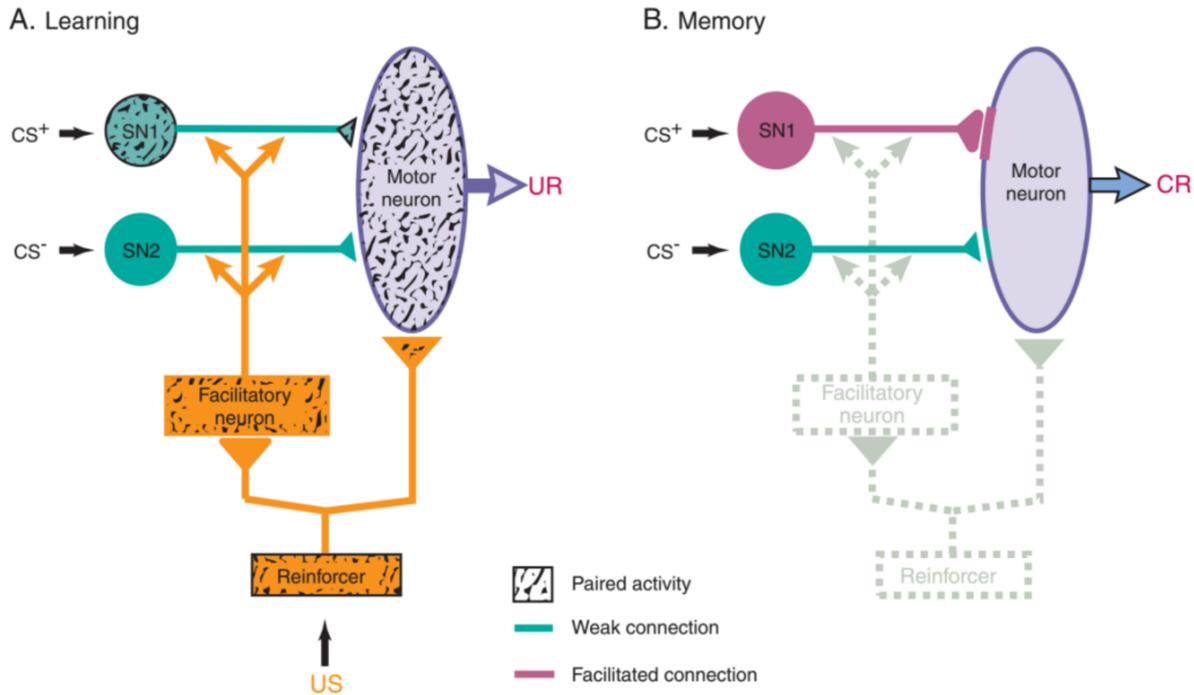
Is through changes in the type and quantity of neurotransmitter that the communication and efficiency between the neuron changes.

This change in the strength of synaptic interaction can be both:

- Functional alterations, as example changes in the amount of neurotransmitter that is released, so it's a short term physiological changes that results in changes in how existing synaptic connections works;
- Anatomical alterations, happen more in the long term, and consist in the formation of the construction of new synapses between neurons;

Hebbian plasticity is a example of plasticity in which already existing synopsis changed the way that they communicate and increase (this is also an example of contiguity learning).

Hebb's Law "neurons that fire together wire together", so if two neurons fire at the same time, the connections between these neuron is gonna get stronger, and they will be wired together.



Example of the circuit that is involved in reflexes in Aplysia.

There is a reinforcer, so in this case is the siphon, and when I touch it, it's gonna activate a neuron that's directly gonna act on the motor neuron of the gill, which closes as response (UR). At the same time touching the siphon also activates another branch of the same neuron, and these activates a facilitatory neuron, which is connected to two different sensory neurons (SN1 and SN2).

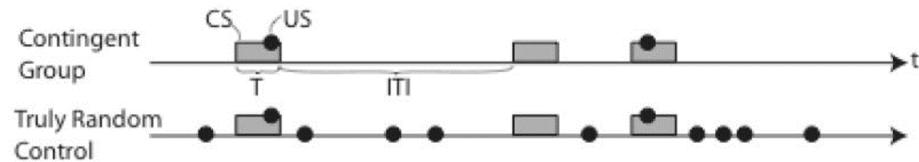
One is the neuron to which I apply this CS class (which is SN1), so the electric shock for example in that case, and this it's gonna stimulate SN1, and another one is a neuron that I'm not stimulating (SN2), which is a control neuron. SN1 is not only stimulated by the CS, but it's also stimulated at the same time by the presence of the US (the yellow arrow), so this is going to fire not only responds to the CS, but at the same time there is also information coming from my US, that also going to strengthen increase the firing.

Instead for SN2, since it wasn't pre activated, is gonna just fire and receive information from the US. After the pairing of the CS and the US on these two neurons if I next activate these neurons, SN1 is going to fire strongly, this connection has been strengthened. So it means that if I apply the CS alone then SN1 is gonna, by itself, activates the motor neuron and the gill response.

6.1.3 Challenging the contiguity assumption

Scientists thought that contiguity was enough, but then in the 60s and 70s evidence came up that quantity is not enough. This was shown in two experimental paradigm one is the **truly random controlled paradigm** and one is the **blocking paradigm**.

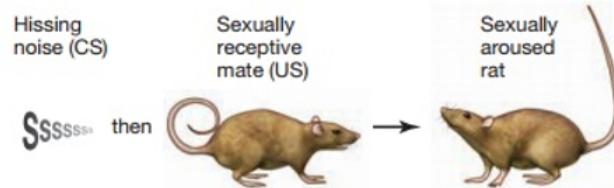
Truly random control experiment in which 2 different groups of people that had performed the same experiment.



And in the two groups the temporal contiguity between CS and US is the same. Also CS are presented together in the same time. In the truly random control group, what happens is that the US is presented sometimes or actually lot of times, also in absence of CS, so it's more likely US is not going to be paired with the CS. Crucially there is no association since the 2 stimuli are not paired.

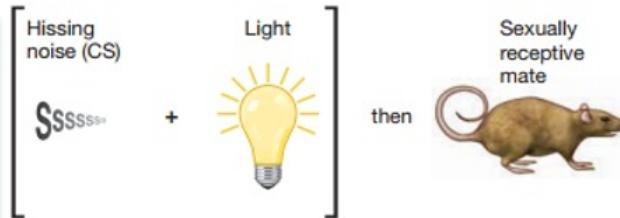
Stage 1

The hiss is reliably followed by the availability of a sexually receptive mate. A CR is thus quickly established.



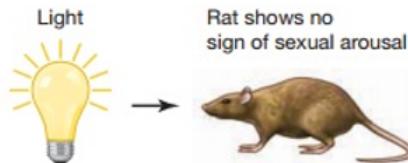
Stage 2

The procedure continues, but now a light turns on at the same time as the hiss. The light is thus reliably followed by the availability of a mate. This seems like a CS (light) followed by a US (available mate), so it should therefore produce conditioning.



Stage 3

Now we see that conditioning has not occurred: The animal doesn't respond (produces no CR) to the light.



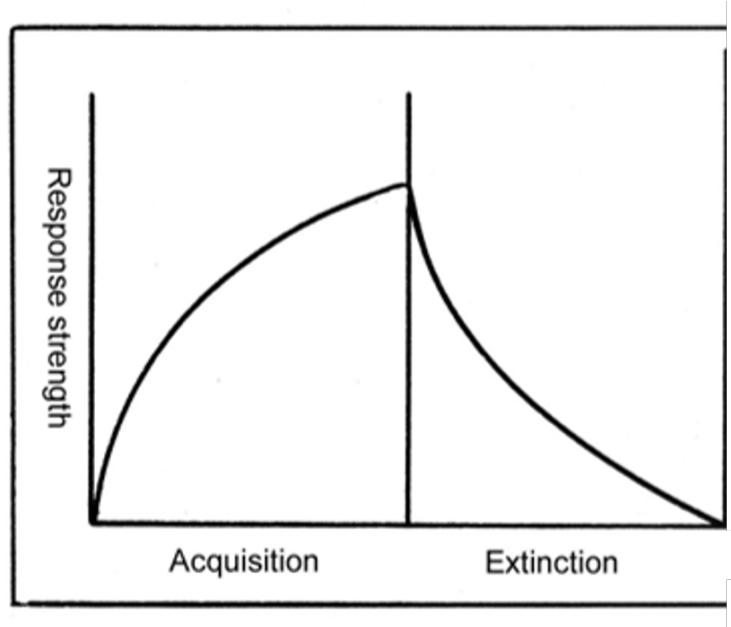
In the blocking experiment it's presented a hissing sound to a rat, and then there is a sexually receptive mate (US). The rat is going to be conditioned and show a conditional response of being sexually aroused, and it also learns that the hissing noise is going to predict the sexually receptive mate. Later the hissing noise it's paired with the light and then it's presented also the sexually receptive mate. Finally it's presented only the light by itself in

order to see if there is the conditioned response or not.

The rat is actually not going to learn, because when the light it's paired with the US and with the noise, the CS doesn't provide any new information about what's going to happen, so there is no need to learn anything about the light. This is called blocking.

6.1.4 Learning from unpredicted event

We **learn from surprising**, unpredicted events. That's also what blocking tells us about. I'm gonna learn something new only if my expectations are violated, so only if there is a difference between what I expect and what actually happens. It's because if my CS or if I add another stimulus, but this doesn't add any new information about what's going to happen next, then there is no surprise, and I'm not going to learn anything about this new stimulus. Expectations can be violated in two different ways, i can expect something and get more or I can expect something and get less of what I expected.



There is an increase in response and in acquisition, as acquisition goes on my condition response is also going to increase until I reach a level in which my expectation fully match what's happening. After few trials, I learn that there is a certain probability that my CS is going to predict the US, and so there is nothing more to learn. And instead in extinction, the opposite happened. At the end of acquisitions the expectation of the US is violated because the experimenters suddenly stops the US. So I expect something that doesn't occur, i get less than what i expected and this leads to a decrease in my condition response.

6.1.5 Prediction error

The difference between what I expected and what actually happened is called **prediction error**.

Prediction error is necessary for associative learning to happen, because if it's 0 then what I expected and what actually happened is the same, so there is no surprise and no learning, on the other hand if the error it's too large prediction did not match at all the observations. So the prediction error has to be sufficiently small.

The prediction error is used to update the value of the condition stimulus. Initially I give a certain value to the reward itself, and then this value is transferred back to the stimulus (cue) that predicts the reward. For example, getting food has a certain value, but then this prediction error is used to update and transfer the value that I give to the food, to the stimuli that actually anticipated or predicted the food.

The cues enable me to anticipate and prepare my response to the outcome that's going to come, so that I can start to prepare for the outcome before it's actually there.

6.2 Computational Model

Psychology started to formalize their learning theory into mathematical models.

6.2.1 Rescorla-Wagner

Rescorla-Wagner model represents the prediction error.

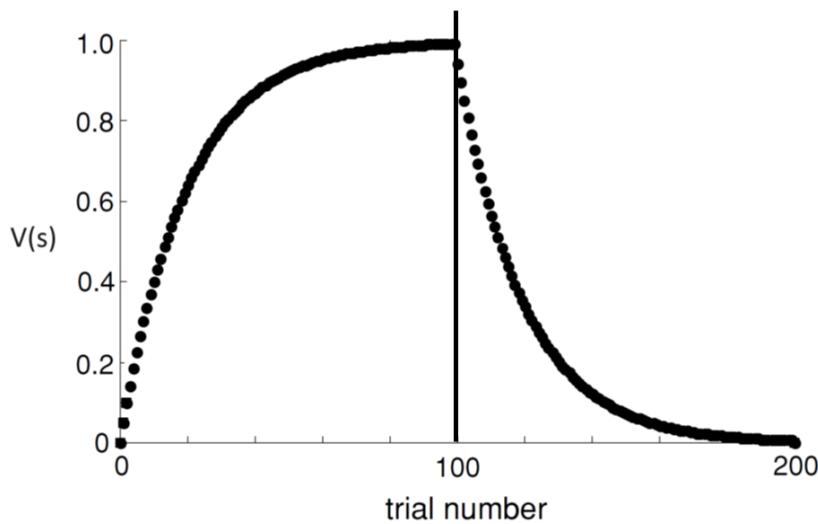
$$\delta_t = R_t - V_t$$

$$V_{t+1} = V_t + \alpha \delta_t$$

R is the actual US at time t, is 1 if I get the reward and 0 if reward is omitted.

V is the expected US at time t.

The expected value V of a give CS at trial t + 1 is equal to the value that I gave to the CS on the previous trials t, plus the prediction error times α , a learning rate parameter.



The shape of this curve is the same as acquisition and extinction. So we can use these models also to model the conditioned response. This model could be used if you want to model the skin conductance response during a condition experiment.

Learning is proportional to the prediction error, so it's larger at the beginning when I don't have a strong expectation about what's going to happen. So there's going to be lots of surprise and then it gets smaller as my expectation becomes more accurate. So the difference between what I expect and what it actually happens becomes smaller.

6.2.2 Temporal difference

Temporal difference model it's a real time model. The major difference from Rescorla-Wagner is that breaks the whole experiment, the trials and what happens between trials in segments of time.

$$\delta_t = R_t + V_t - V_{t-1}$$

$$V_{t+1} = V_t + \alpha \delta_t$$

The equation is basically the same, but in this model it's considered only a moment in time.

6.3 Dopamine

A neuron can translate positive or negative prediction errors in a signal by changing the amount of firing. Increasing their firing rates when the error is positive and decreasing when the error is negative. These changes in firing rates can be achieved through synaptic plasticity. So changing in the efficacy of the signups, by changing for example, the amount of

neurotransmitter that is released, or also **neuromodulator**.

Dopamine it's a neurotransmitter that encodes the prediction errors in the brain. It plays many roles in the nervous system and in behavior, for example is important for motivation, learning, decision making, reward processing.

6.3.1 Dopaminergic pathways

There are 3 major pathways of dopamine

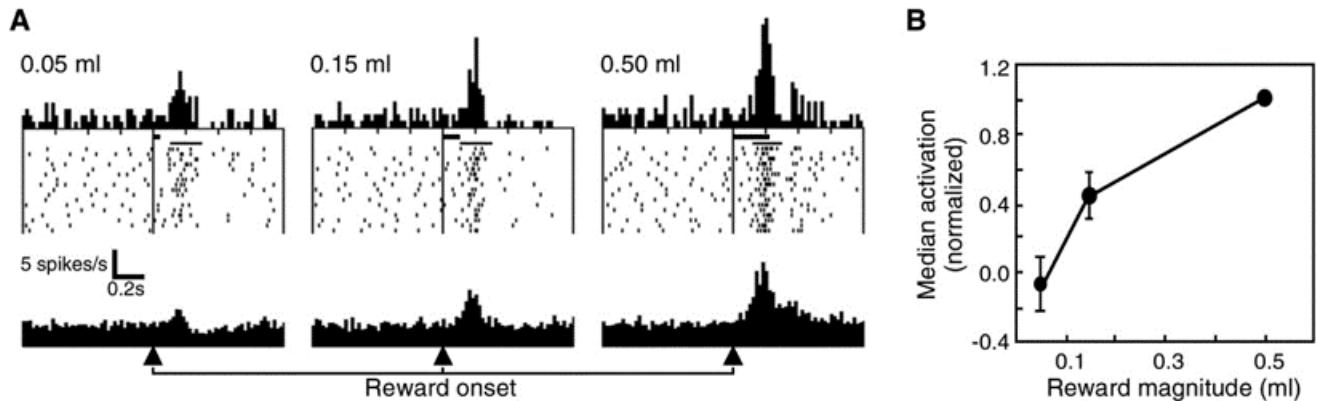
1. Nigrostriatal pathway: It starts in the substantia nigra and it's critical for movement, since it's part of the basal ganglia motor loop. If some part of this pathway is depleted of dopamine, then you're going to have either Parkinson or Huntington. For example, the loss of dopaminergic neurons in the substantia nigra produces over activation of the indirect pathway that leads to that decreases movement and leads to Parkinson, and the opposite happens for Huntington's disease;
2. Mesolimbic pathway and Mesocortical pathway both originates in the VTA and are crucial for reinforcement learning;

6.3.2 Hypothesis dopamine neuron functioning - Animal

In a paper published in 1954 it's described the effect that electrical stimulation on certain areas of the brain had on the behavior of rats. Olds and Milner realized that the area where they put the electrodes, were areas where there were dopaminergic neurons. So they were stimulating the release of dopamine in the brain. There was an extremely rewarding effect of the stimulation in the animal, and then this reward was controlled, so the animals started to produce the behavior again and again in a compulsive way because it was obtaining a reward because dopamine was released in the brain.

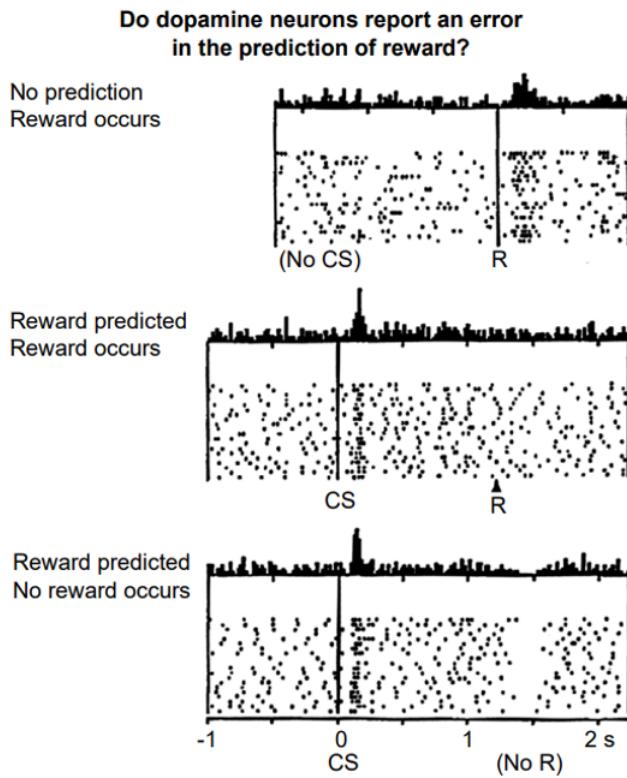
This evidence about dopamine leads to the discovery that it's a **broadcaster of reward signal** (this discovery was later denied when it was discovered that dopamine actually carries the prediction error). Wolfram Schultz conducted a research on monkeys and he recorded directly the electrical activity of dopamine neurons. He put electrodes in monkeys brain and measured the difference between the intracellular and extracellular electrical potential. Inspecting this measure is possible to deduce if the neurons are firing.

Schultz realized that actually these neurons weren't firing when the movement started or when the monkey was executing the movement, but they fired when the monkey was touching the reward, and only if this reward was unpredicted.



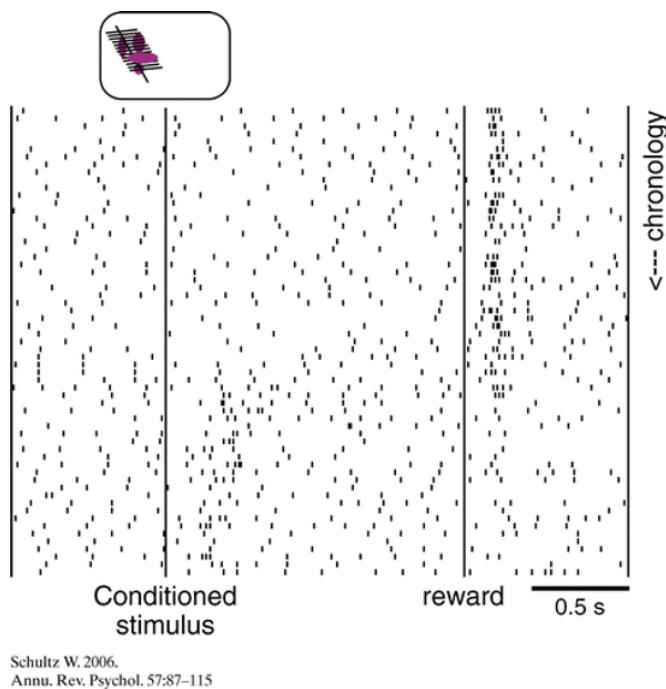
These measures in ml tell you how much juice was given to the monkey, and the firing rate of the neurons changes depending on the amount of the reward, because the dopamine is able to distinguish between rewards and also how much reward you got (**magnitude of the reward**).

The same amount of reward can elicit two different types of activity in the dopamine neurons, because they process reward magnitude relative to a predicted magnitude, so if one stimulus can predict both small or medium volume of juice with equal probability, there will be actually 2 possible outcome. The reward is objectively the same, but the neuronal response to the reward is different.



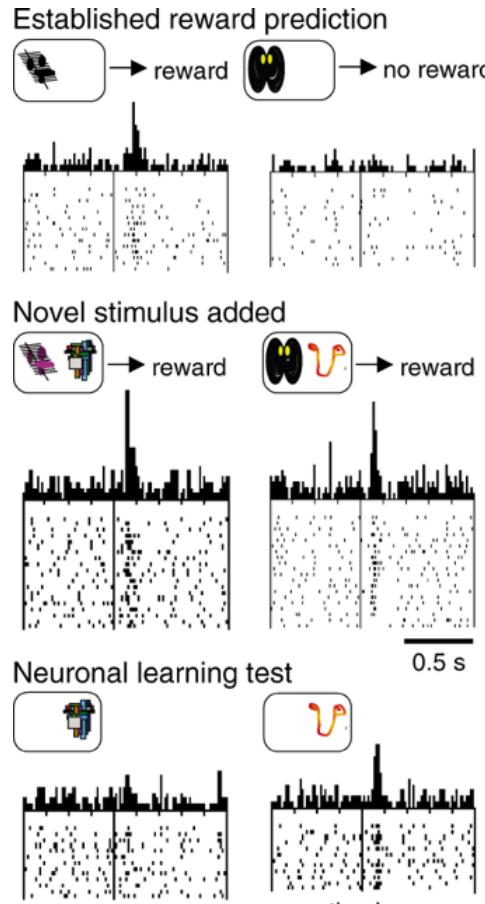
In these 3 different scenario there is time on the X axis, and every black dot represent a neuron that's firing:

1. In the first, appetitive juice is given to the monkey without any CS, hence there is a **positive error** in the prediction of reward, and the dopamine neuron is activated by this unpredicted occurrence of juice;
2. In the second, first the cs is paired with the reward, and the reward occurs according to the prediction, hence there is **no prediction error**. The dopamine neuron is activated by the reward predicting stimulus but fails to be activated by the predicted reward;
3. In the last scenario, after the CS there is a peak, it means that the animal is indeed expecting a reward. This prediction is violated because I don't give the rewards. This should produce a **negative prediction error** which should decrease the firing rate of the neurons, and this is indeed what happens, and this doesn't happen at a random time, but exactly when the animal expect the reward. Somehow the dopamine neurons have understood that this is the time when they should get their reward;



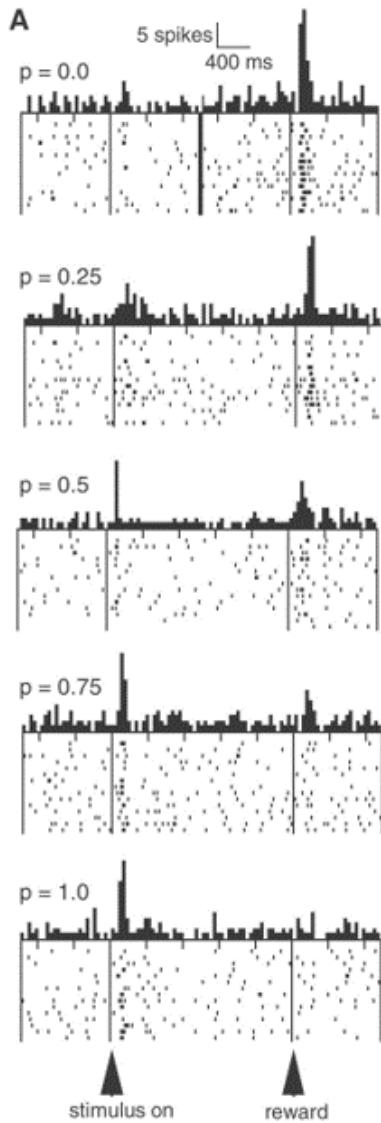
When associative learning happens, the value that first given to the reward is given now to the CS. Here you are looking at what happens during training over time. Experimental trials are represented on the Y axis. At the beginning there is a lots of neurons firing not in response to the conditioned stimulus, but in response to the reward, because the monkey doesn't know what to expect in this case, it sees the visual stimulus and doesn't know what it means. Then it gets a reward, so there is an increase in dopamine firing (dopamine release).

And then after several trials, the monkey paired the picture with the drop of juice, so the neural firing increases not anymore to the reward, but actually increase when it's presented the condition stimulus. What's interesting is that there is some overlap when the monkey is learning. There is some time when the monkey is making the association, when there is a increase in dopamine firing both to the condition stimulus and to the reward



Schultz W. 2006.
Annu. Rev. Psychol. 57:87–115

Neural learning is **blocked** when the reward is predicted by another stimulus, because the neuron has the capacity to respond to reward-predicting stimuli (top left) and discriminates against unrewarded stimuli (top right). The addition of a second stimulus, results in maintenance and acquisition of response, respectively (middle). Testing the added stimulus reveals absence of learning when the reward is already predicted by a previously conditioned stimulus (bottom left).



Tonic neurons are never silent for a long time, they have a basal activity firing rate. In the case of dopamine neurons, you always have a certain amount of dopamine floating around your brain and your synapses and that is your base level of dopamine. It means that the neuron has to fire with a certain frequency and this is the tonic activity.

Phasic activity consists in a sudden silencing of the release of the openings, on silencing of the dopamine neurons. At that happens when there is a negative prediction error, for example you were expecting a reward, but you didn't get any.

Phasic activation of dopamine neurons vary monotonically with reward probability.

With a low P of receiving reward after CS, the dopamine neurons firing is high when you expect to get the reward. With a high P of receiving reward after CS, the dopamine neurons firing is high right after the CS. With a P of 0.5 to receive the reward after CS (scenario with max uncertainty) there isn't high peaks neither after CS nor after receiving the reward.

Scientists also realized that also **tonic** firing of dopamine neurons **encodes the probability** or the **uncertainty** that you're going to get a reward, and in particular they encode these when the uncertainty is maximum (50%).

Dopamine neurons also **encode timing** of when the reward should arrive. In particular when the reward is given at an unexpected time, later or before the timing, there is an increase in dopamine neurons firing.

Also the dopamine neurons **stop firing** when no reward is given, even when there is no stimulus in the environment, and this doesn't constitute a neuronal response to a stimulus but reflects an **unfulfilled expectation**.

Dopamine neurons

- DO NOT broadcast a reward signal
- DO broadcast a prediction signal

Summary

Dopaminergic neurons exhibit a strong phasic response to:

- an unexpected reward
 - discriminating between reward no-reward
 - discriminating between reward magnitude (more or less reward)
 - * in a relative way, rather than absolute
 - discriminating between reward probability (more or less likely to get reward)
 - timing (earlier or later reward)
 - in a signed manner
 - * increase firing to unexpected delivery (positive PE)
 - * pause firing to unexpected omission (negative PE)
- transfers back to a cue which predicts reward occurrence (i.e. CS)
 - enabling associative (reinforcement) learning

6.3.3 Hypothesis dopamine neuron functioning - Human

Studies in humans have shown that there is a reward prediction errors in the human brain.

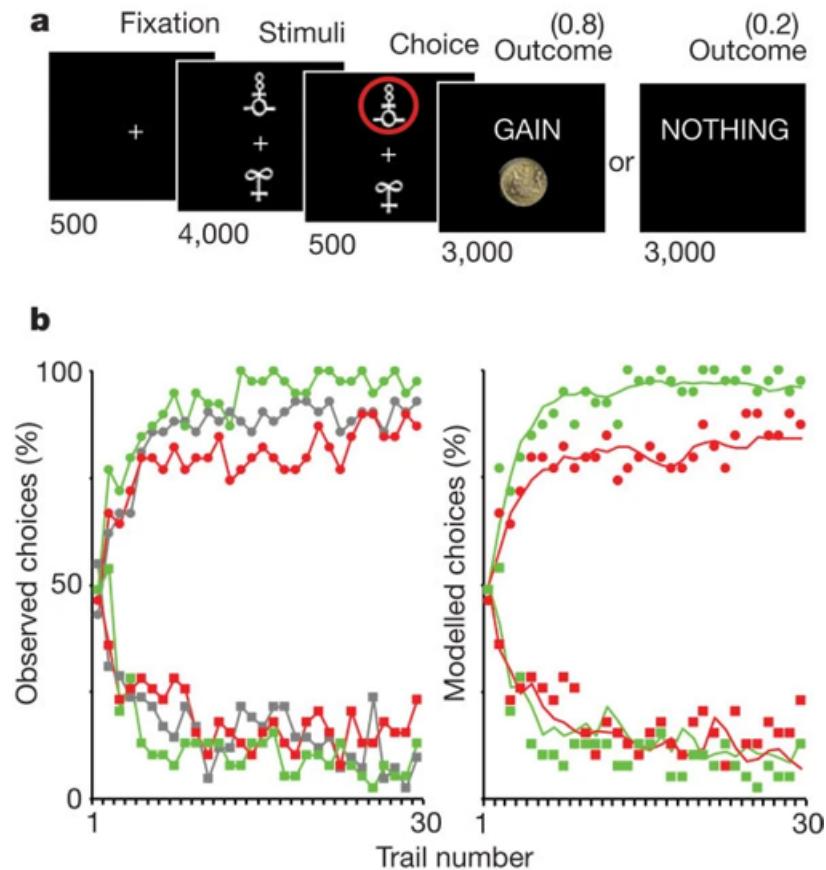
If we have an unexpected win there is an increase in the firing rate relative to expected win, and relatively you decrease the firing rate for the unexpected loss.

In an experiment scientist designed a fixed sequence of reward, which was composed by the

repetition of water-juice and the delay. The time delay between the water and the juice was always the same, so for the human brain it's predictable what you're gonna get and also how much time is gonna pass between the two stimulus. The scientist also replicated the experiment without a fixed sequence of reward and delay, in order to measure the difference in brain responds between the unpredictable and predictable conditions. The scientist measured the activity in the brain, in the two different type of experiments. They subtract activity in the unpredictable experiment from the predictable experiment and so they found the areas (nucleus accumbens and bilateral superior parietal cortex) that was active only in the unpredictable experiment. They performed the opposite operation for the predictable experiment.

The scientists found that the **predictability in time** modulates the human brain responds to reward.

Scientists recruited participants and they gave them drugs, different drugs that were **agonist** or **antagonist** of dopamine, like **haloperidol**, and what it does is that it can mimics the chemical structure of dopamine and it can binds to the same receptors, but they don't have any effect, so they **block** the effect of the **receptor**. The scientists were seeing that they gave a certain task to participants, and the performance on the task depended on whether they gave to participant the agonist or the antagonist.



Participant had two different symbols appearing on the screen, and they were asked to choose one or the other. And in 80% of the case, for example, one symbol led to gain of a certain amount of money, or in 20% of the case, instead, gain nothing. They have to learn which symbols is going to give them the reward. This is **instrumental learning** because there is an association between behavior and outcome.

What you can conclude is that on the top, dopamine seems to indeed have a role in behavior in these games because if the dopamine is increased the participant is going to learn better wrt participant that has less dopamine.

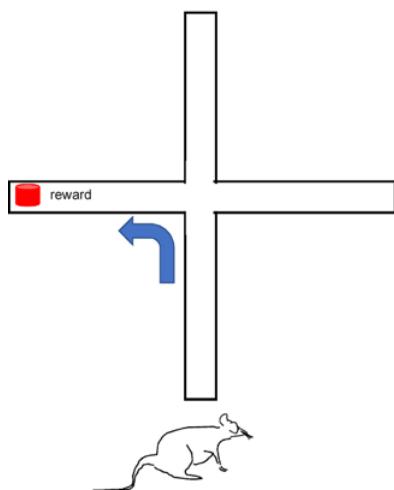
But when you have to learn about avoiding a loss the dopamine doesn't seem to have such crucial role, because there isn't really a difference in performance between the different groups. So regardless if participants taken L-DOPA or other dopamine antagonist, they don't change their performance.

7 Model Free and Model Based Learning

7.1 Learning strategies: map vs response strategy

There are different learning strategies for navigating, in particular there is 2 types of strategies, one is the **place/map strategy** and the other one is the **response strategy**. In the response strategy, you just learn a particular sequence of motor responses, so for example that turning left is gonna get you out of the station, instead with the cognitive map strategy, you try to learn the whole map of the place where you are.

1.



The rat during this experiment has to learn to navigate a simple maze in order to get out of it and get the reward. Scientists put the rat at the bottom starting point, and he turned left at the intersection, finding the reward. Did the rat use the response strategy or the map strategy? To figure out the scientists placed the rat in the top starting point of the maze, and observed which turn he takes at first cross, if the rat took the left turn, it means that he learned by a response strategy (turning left = reward). Instead in this case it's wrong because the rat started from a different point.

If the rat had learned with the place strategy instead it would have learned the whole layout of the labyrinth, and he would know where to turn.

Psychologist began to study these 2 types of learning strategy and from these they arrive to what's called **model-free** and **model based** learning or **goal directed** and **habitual learning**.

7.2 Generations of studies about goal directed and habitual learning

When you do something to obtain a reward or to avoid punishment, you are performing an action. How do we decide which is the appropriate action to take? Are we flexible in taking this action or do we are directed by the goal we want to achieve, or even we act based on our past experience?

Ray J. Dolan and Peter Dayan reviewed all the literature about goal directed and habitual learning and they identified different generations of studies:

- Generation 0, cognitive maps vs stimulus-response [experimental psychology]
- Generation 1, goal-directed vs habitual actions [experimental psychology]
- Generation 2, goal-directed vs habitual actions in the human brain [cognitive neuroscience]
- Generation 3, model-based vs model-free computational analyses [computational neuroscience]

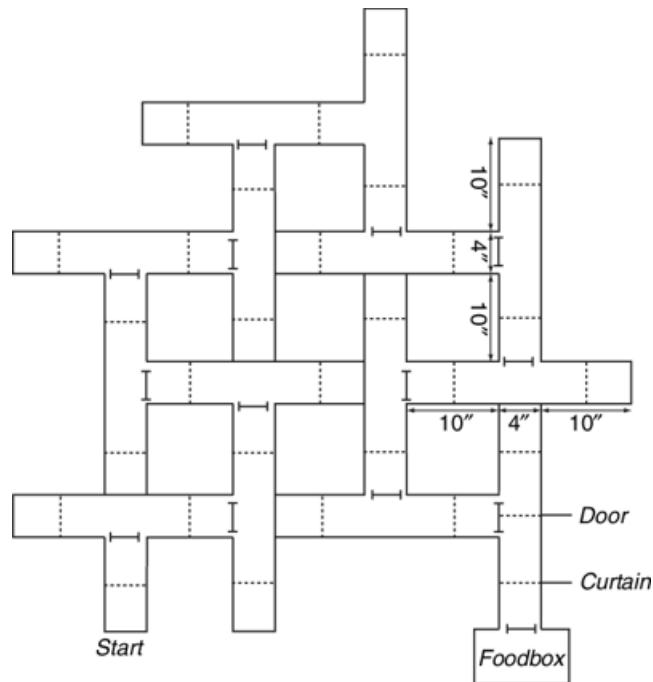
7.2.1 Generation 0: cognitive maps vs stimulus-response

Before **generation 0** there was a big debate on how the rats solved the maze, and the prevailing view was the stimulus response theories, which originate from associative learning and behaviorist approach (we just study behavior, we don't study what's in the black box).

Stimulus response theory is about a one-to-one connection between a stimulus and a response, and the learning depends on the strengthening of the stimulus response connection. Furthermore, according to the stimulus response theory, the animal is helpless in responding, it automatically responds to that stimulus without any control.

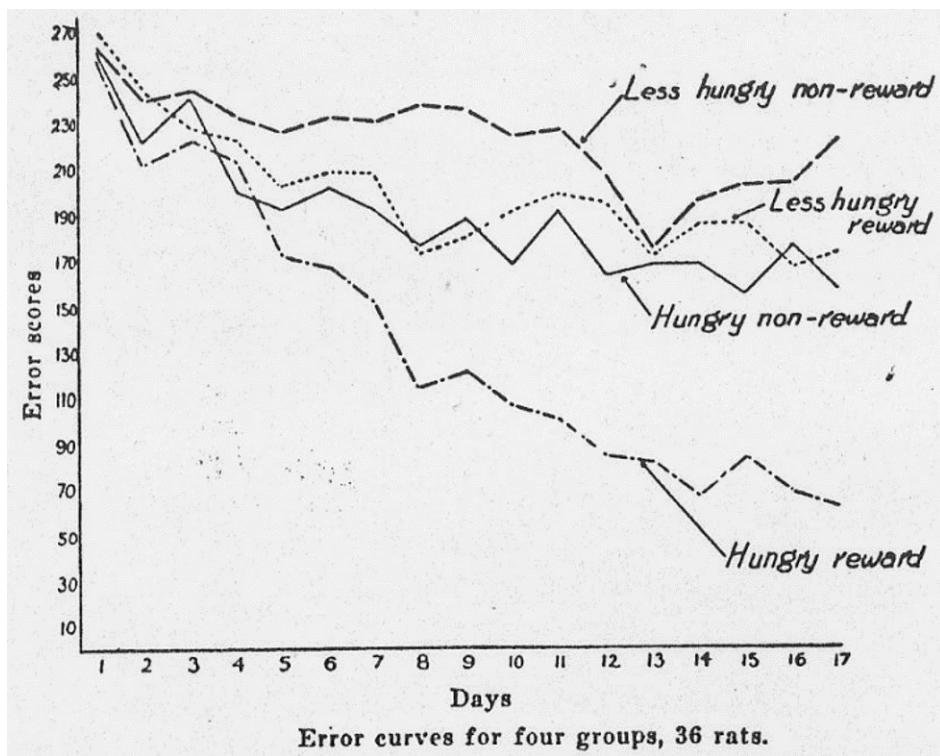
In the **Field theories** solving the maze means that we have to learn about the actions that are gonna get us the reward. So solving the maze is a matter of creating a mental map of all the connections that are gonna lead us to the reward. So if you start in another position of the maze, you're just not going to repeat the sequence that you've learned in the beginning, but you adapt the sequence of actions depending on where you are in the maze, by the means of the mental map.

To see what theory was the best fitting or which theory corresponded to how rats actually learned to navigate the maze the scientist Tolman conduct lots of experiments using a maze created by him, which is a simple maze with the addition of door and curtain, so that the mouse could not see what was next at a crossing. At each step the view of the road is blocked, so he has to decide where to go, and only after he realizes the consequence of that action.



Tolman put 2 groups of rats in the maze

- Group 1: no reward for solving the maze;
- Group 2: group of hungry rats that get food for solving the maze;



The group 2 was the faster to solve the maze.

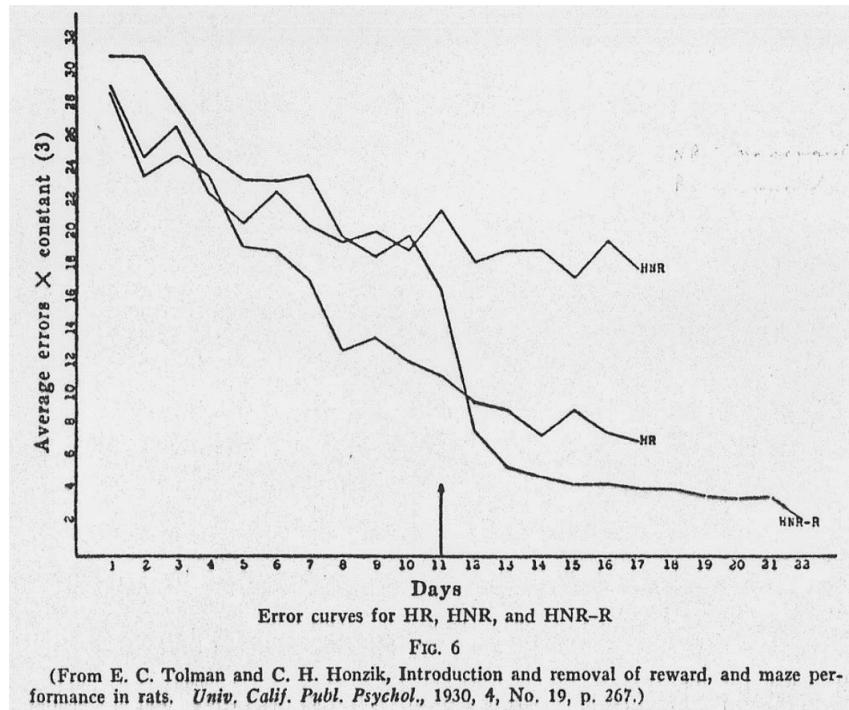
Y axis --> number of error that the rats made before arriving at the end of the maze.

And in particular as the day goes, hungry rats that were rewarded were faster in learning, so decreased the amount of error compared to the other groups. The 1st group needed more time because they had less motivation to finish the maze, so motivation seems to put a big role in what you learn.

Toleman did another experiment to test if motivation really plays a role, according to stimulus response theories. This time he had 3 group:

1. Group 1: no reward for solving the maze;
2. Group 2: food reward for solving the maze;
3. Group 3: food reward for solving the maze only occurs after 10 days;

According to stimulus response theories, in the 3th group no learning should occur, until they get the food.



Group 3 performance at day 11 **improve dramatically** and they even get better than the rats that always got the food.

Seems that the rats knew all along how to navigate the maze, but the improvement in performance showed that they've learned only when they got the food. So indeed, they've learned something even when they were not rewarded, they indeed learn how to navigate the maze, it's just that they show what they've learned only when they are motivated to show it. So this is in contrast with stimulus response theory.

This evidences tells us that you always learn even when you are not rewarded, this it's called **latent learning**.

What's important of generation 0 studies it's that it challenged the stimulus response theories, because it says that we don't only learn about stimulus and responses, which correspond to model free learning, but we also learn cognitive maps which correspond to model based learning.

And it also challenged behaviorism because Tellmann showed that there is also latent learning, so you can still learn, but you are not gonna show it until I motivate you to show it.

It doesn't mean that the stimulus response theory is wrong, but we have both type of strategies from which we can learn.

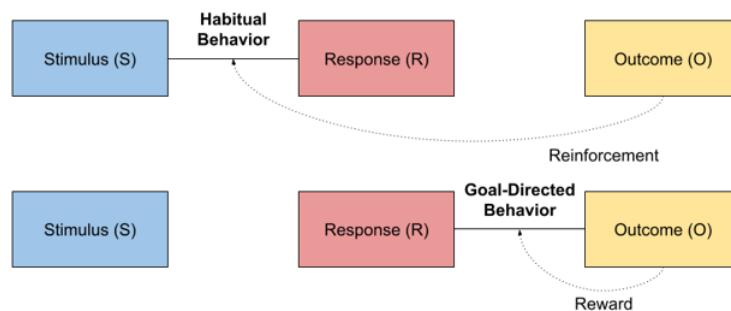
7.2.2 Generation 1: Goal-Directed vs Habitual actions

Generation 1 studies focus on goal directed and habitual behavior/actions.

Researchers started to wonder if cognitive maps or stimulus response strategies could be used also in learning about nonspatial task.

Cognitive map -->goal directed action. newline Stimulus response -->virtual action.

In generation one we are still talking about the relationship between stimulus, response and the outcome. The difference between whether you decide to perform an action or another could be because the action performed it's an habitual behavior, or a goal directed behavior. When we talk about habitual behavior, the relation between the stimulus and the response is strengthened, and the relationship between the response and the outcome becomes less important, it's less strengthened.



Habitual behavior it's an action that you do because it's a habit you have to do it, it's an action triggered by a specific stimulus.

Instead when the behavior is goal directed, even if you see the stimulus, you're gonna do the response or not depending on your goal.

There are two criteria that define our behavior as goal directed:

1. You **know the relationship** between the response and the outcome;
2. You are **motivated** to get the reward, as example getting food if you are hungry;

Goal directed behavior, involves active deliberation but it has a **high computational cost** wrt habitual behavior, because you have to keep track of outputs of all actions. In addition, goal directed behavior is **flexible**, it's more adaptive because when you don't want the food anymore (no motivation) or there is no more food, you're not gonna spend energy in performing the action.

Habitual behaviors (compulsive) is when you perform an action automatically just because it has been rewarded in past. What's important is that the action it's not influenced by the current value of the outcome. So even when the outcome doesn't have any value for you anymore or any motivational value, you still do the action and even not only when you don't need it anymore, but even when the outcome is become aversive. So even if you don't like it anymore, you're still triggered to do the action.

There is no deliberation about whether to do the action or not. The response is **inflexible** to changes in the environment and has a **low computational cost**.

For testing if a behavior is goal-directed, we first do the training, in which the rat associate that he has to press the lever to get the food.

Then afterwards, we give the animal lots of food until it doesn't want food anymore, this process is called **devaluation procedure**, crucially we remove the motivational value of the stimulus, in particular, in the goal directed behavior we are taking out the rewarding value of the outcome, and this should stop the behavior.

If the rat stop to push the lever then him behavior was goal-directed otherwise was habitual.

Another manipulation, other than devaluation, is **contingency degradation**. We teach to the animal that pushing the lever doesn't lead to the food anymore.

Suppose a different scenario, where I overtrain the rat to press the lever for days and days. After this train the mouse still keep compulsively pushing the lever even if he doesn't want to eat anymore, so with overtraining, I've turned a behavior that first was goal directed into an habitual behavior.

1st generation studies also looked at the neural basis of this distinction between the goal directed and habitual behavior, and they found that the striatum was important for these two types of behavior, and in particular, the dorsomedial part was important for goal directed behavior and the dorsolateral for the habitual behavior.

Striatum it's important for prediction errors or reward prediction errors and for goal and habitual actions. Striatum receives inputs about prediction errors and turns those prediction

errors to action.

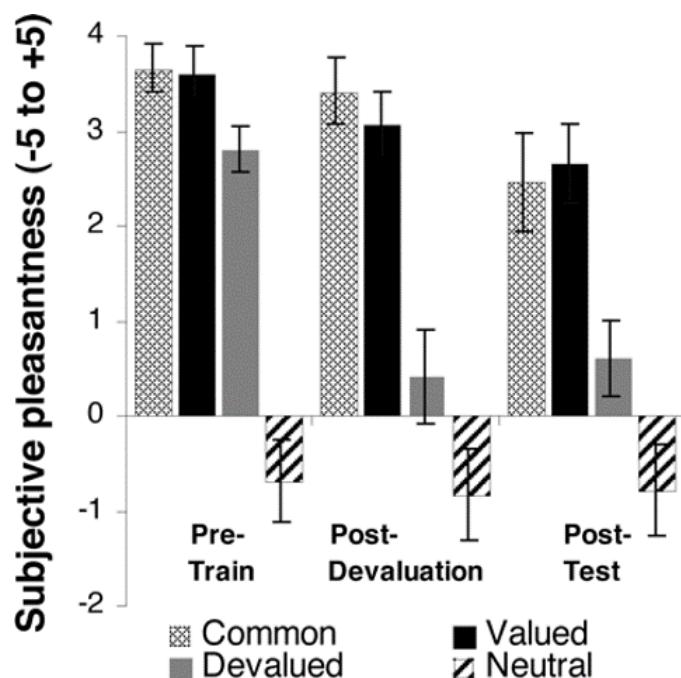
So it's an important structure that links together prediction errors, motivation and learning and turns it into action.

7.2.3 Generation 2: Actions and Habits in the Human Brain

In **generation 2** scientists wondered if the evidence that was found in rats could also be observed in the human brain. So they translated experiments from animals to humans and used fMRI (Functional magnetic resonance imaging) to investigate whether also in the human brain, there was difference between goal directed actions and a habitual actions.

In this experiment what we have is that you see a screen, and every time you see two different pictures, and you have to choose one. Depending on which image you choose you get different outcome, such as chocolate and orange juice. You need to make sure that whatever you get, it's because of your main experimental manipulation, so in the first learning stage you have to learn the association between different images and corresponding outcome. And you should also learn to choose more the pictures that are gonna give you the reward with higher probability.

Later, as done with the animal, scientists give you lots of tomato juice until you don't want it anymore, and then you have to choose again between these pictures. You should now decrease the choices that leads to tomato juice and instead increase the choices of the chocolate if your behavior is goal directed, if this doesn't happen then you behavior it's habitual



Before and after the training was asked to participant to give a score of pleasure of the

juice and chocolate, looking at the result there are evidence that the food that was repeatedly given to humans (devaluated) received a drastic drop in votes.

The goal directed behavior is related to activation in different parts of the brain, in particular the **orbitofrontal cortex**.

In this experiment we had two group of participants. One group of participants had extensive training, it means that it does the task six times more than the other group of participants. The other group did the task only 1 time.

Task: a fractal image was shown on the screen, and the participant has to press the button indicated in the image, among the 4 buttons, if he pressed the right one, after each press, either a grey circle briefly appears (50ms), indicating no reward, or a picture of an MM or Frito appeared, indicating a food reward corresponding to the picture. Participants could press the button as many times as they wanted so they could get candy as many times as they wanted. They learn the different images of fractals and the button associated.

At the beginning the two groups are the same, they just press at a certain frequency, which is similar between the groups and regardless of the reward that you're gonna get. Then afterwards I overtrain and also I devalue one of the two, therefore the overtrained group (3 days of training) changed their behavior from directed to habitual.

Such that even though they don't want the outcome anymore, they are still going to press button for it. Instead with the one day training group, the behavior is called directed, so they only perform the behavior that is gonna lead to the outcome that they actually want.

The evidences gathered during the experiment shows that different parts of the striatum were important for both directed and habitual behavior.

7.2.4 Generation 3: model-based vs model-free computational analyses

In **Generation 3** studies we have the development of the computational formalization of the goal directed and habitual behavior, from an algorithmic point of view.

Computationally, when we referred to goal directed actions, we refer to model based algorithm and instead habitual actions are described by model free algorithm. OK and what's important is that when we use the word model in this case.

Model free means free of a model of the world (environmental contingencies), and **model based** instead is an algorithm that includes a full representation of the state of the world.

In the model free each state is separate from each other and you only compute whether that state is gonna lead you to a reward or not, while in the three system each level of the tree is connected to the other ones.

So you don't only have to compute the outcome of an action, but also keep track if that action in that specific state gave you the reward or not, and how that is going to affect the

future states, of course this is more computationally costly.

A specific type of sequential of Markov decision tasks, a 2 steps Markov decision task, where the second step depends on the action that you're going to make in the 1st state.

Each of those action has a probability to lead to a different second stage and then the choice that you make at the second stage has another probability of leading to another stage. Thanks to this type of task, it was possible to distinguish if the agent was learning and performing the actions in a model free or model based way.

While you're performing the task you should learn the connection between picture and state, and between state and reward, in order to maximize the reward.

In general **model free agents** prefer to **repeat** actions that lead to reward **regardless of the likelihood** of that transition.

If you are **model based** instead what happens is that you prefer to perform an action that is **commonly rewarded**. The opposite happens with punishment, if you got punished from a common transition, you're gonna avoid that picture and choose the rare one, because even if it could leads to punish too, it's more rarely that happen.

Participant were tested in order to check if they had learned something from the first session, even though no rewards were given to participants. By looking at their choices, if they've latent learning they are gonna perform well on this task and choose the stimuli that actually gonna lead to the reward.

This type of learning is possible only with model based but not with model free learning, because model free learning depends only on reinforced actions (no reward \rightarrow no reinforcement). (Model free learning focuses only on predicting rewards without building a model of the whole environment. So if there is no reward there is nothing to predict and there is nothing to learn about.) The results was that 13 subjects out of 18 made the choice that reflected latent learning, so it means that they indeed have learned something.

After that scientists tracked the whole sequence of behavior that was made during the whole session and fit the data into computational models, in order to see if their behavior was determined by model free algorithm or model based, and they actually found that what explained best the model was neither model free or model based, but was on a **hybrid model**.

The model free algorithm has a **reward prediction error** (similar to Rescorla-Wagner), while the model based algorithm encode the **state prediction error**.

They looked at neural data to see what kind of brain areas were involved in the two different types of learning. It was noticed that the reward prediction error was encoded by the **ventral striatum**, while the state prediction error was found inside the **lateral prefrontal cortex**.

The researchers were looking to find both model free and model based **signature** in the

brain **at the same time** when the learning was both model free and model based at the same time.

The main structure that's involved in the coding of prediction errors doesn't code only reward prediction errors, but also state prediction errors.

Generation 3 results challenge the notion of a separate model-based vs model-free learner and suggest a more **integrated computational and neural architecture**.

In the brain, there is a dynamic **inter-dependency** between goal-directed/model-based and habitual/model-free systems, which may act simultaneously and competitively.

References

- [1] Rodney Brooks et al. “Is the brain a good model for machine intelligence?” In: *Nature* 482.7386 (2012), p. 462.
- [2] *SeveredCorpusCallosum*. URL: https://www.youtube.com/watch?v=lfGwsAdS9Dc&ab_channel=ctshad.