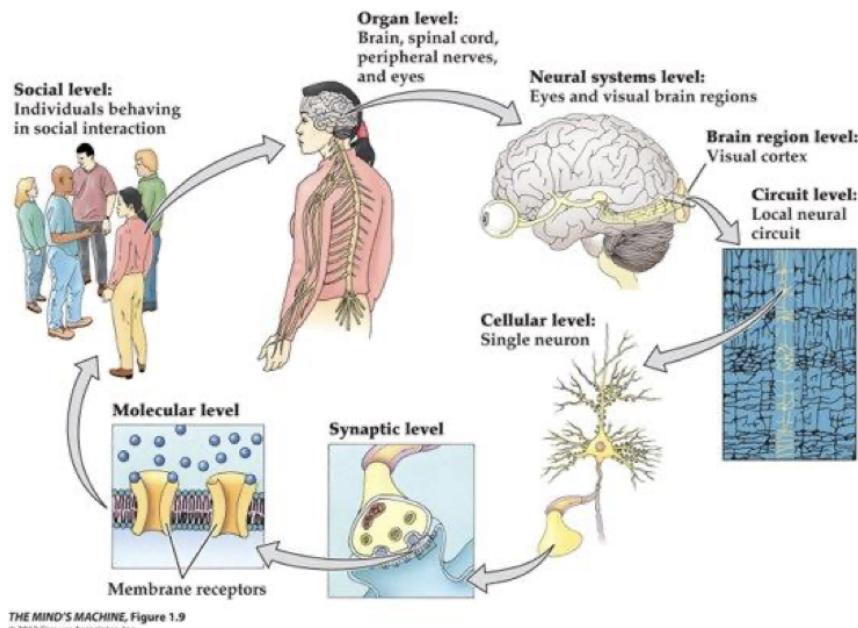


Module 1

What is Cognitive Neuroscience?

@February 22, 2023

Neuroscience → is about the structure. Study of Nervous System organization and functioning on a chemical/cellular level, transmission of information and communication between cells (micro-level). It also deals with the macro-level, which is how communication between neurons and different systems works.



Neuroscience

It is a multidisciplinary science, comprising Physiology, Anatomy, Molecular Biology, Cytology, Computer Science and Modelling. Also includes different levels of investigations, from molecular (neuroanatomy, signaling) to circuital/systemic (perception and multi-sensory integration, reflexes and actions).

Cognition → many different types of processes. Deals mainly with functions, more correctly mental processes, so learning, attention, memory, perception and thoughts. Deals with the “abstract” functioning. Relates to acquisition, storage, manipulation and retrieval of information (human thought processing).

- Perception → take in information from the environment through the senses

- Attention → allows to focus on a specific stimulus in the environment
- Learning → manipulating new information and integrating it with prior knowledge
- Memory → encode, store and retrieve information. Critical component in the learning process
- Action → use perceived information to interact with the environment
- Language → the ability to understand and express thoughts through spoken and written words. It allows communication with others
- Thought/Higher reasoning → allows to engage in decision-making and problem-solving

Cognitive Neuroscience is therefore the study of how the different cognitive functions emerge from the nervous system (how structure rises to function).

Why it is a good idea to use the brain as a model for AI? The human brain is the existing proof that general intelligence is possible, so by emulating this system we can replicate the idea. Plus, studying animal cognition and its neural implementation can provide a window into different aspects of higher-level general intelligence.

Functional similarities → it is a all-or-none firing system, which is similar to binary computations (though this can't capture the complexity of the brain).

Neuroscience can be a source of inspiration for new types of algorithms and architectures, independent of and complementary to the mathematical and logic-based methods and ideas that have largely dominated traditional approaches to AI. Plus they can also provide validation of AI techniques that already exist: try the algorithm and see if the nature works similarly: positive outcomes give strong support for its plausibility as an integral component of an overall general intelligence system.

Different fields try to create technology inspired by nature (biomimicry) → Biomimicry is the emulation of models, systems and elements of nature for the purpose of solving complex human problems. Living organisms have evolved well-adapted structures and materials over geological time through natural selection.

The animal brain, though, is not necessarily a good model for AI → The engineer perspective of “If it works, it's fine”. We still don't know the exact functioning of the brain, so it's still kinda limiting. Plus the human brain does not implement intelligence the same way as a computer: computer does not excel in the task of generalization as humans do provided with the same amount of data (they require a rather huge collection of examples). Plus in the end it would become a circular approach (might exclude us from evolution in modeling).

Hardware and software are separate in machines, but intimately integrated in the human brain. Living cells generate both electrical and biochemical changes. Plus they are more akin

to a Recurrent Neural Network (they operate in circles) rather than a Feed-Forward Network (linear chains of causality).

There are different levels of brain emulation. Different companies are trying to emulate at different level. E.g. structural emulation → replicate the structure to obtain similar functions. Blue Brain Project. A more abstract level is to mimic the function, how the brain works, and be inspired to gain insight on the general mechanisms.

Brain Emulation:

Structure → closely mimic or directly reverse engineer the specifics of neural circuits (Blue Brain Project)

Function → mimic the computational and algorithmic levels of neural systems, to gain transferrable insights into general mechanisms of brain function (Deep Mind)

What is Cognitive Neuroscience?

In the 70s in New York, in a Taxi going to the Algonquin Hotel, there were Michael S. Gazzaniga (expert on Split Brain studies - connection between the hemispheres is severed by eliminating the functionalities provided by the corpus callosum which can be a solution to severe epilepsy cases) and George A. Miller, they wanted to study how the structure of the nervous system would give rise to cognitive processes. The name stuck.

NB → the different hemispheres are highly integrated. Language functions reside mostly on the Left Hemisphere.

Historical Perspective

Neuroscience Side

Oxford, 1650. Anne Green is sentenced to death for a crime she didn't commit (stillbirth). Thomas Willis performs an autopsy of the woman; surprisingly, she was still alive. He actually made money out of this, and was the first to link abnormal behaviors to changes in brain structure, foreshadowing cognitive neuroscience with the notion that isolated brain damage (neurology) could affect behavior (psychology). Pioneering research into the anatomy of brain, nervous system and muscles.

Cue 100 years. Franz Joseph Gall theorized that the brain was the organ of the mind and different processes were localized in different regions of the cortex (father of Localizationism). If the person used one faculty with greater frequency, the specific part would grow, causing a bump in the skull (he went too far with this). So by analyzing the skull dents, you could read the personality of the person (Phrenology). Rather popular idea

actually. He wasn't trying to disprove hypotheses but just to confirm them, so he actually wasn't a scientist in this sense. Napoleon called to have his skull analyzed. He didn't receive the news he hoped and asked Flourens to disprove the theory, which he did. But for more advanced abilities (memory or cognition) he couldn't find specific areas, so the whole brain had to participate (Aggregate Field Theory).

This leads to two different sides (we indeed have both, we have certain areas for certain functions but we need the whole system to actually work):

In favor of Localizationism:

Dax reported 3 autopsies on people with speech impairment and found similar lesions to the left hemisphere. **Jackson** studied epileptic seizures (huge uncontrolled increase in electrical firing of neurons, often involving motor cortex) and found that it progressed in an orderly manner from one part of the body to another. Hypothesis: different parts of the body are represented in the motor cortex and could propose a topographic organization of the cortex (this can be replicated also for other areas, e.g. vision, sensory systems). **Broca** individuated a lesion in the left hemisphere in the inferior frontal lobe of a patient called Tan (he could only say this word). This is Expressive Aphasia, impairment of language production, but untouched comprehension both of the language and of the inability to communicate. Also **Wernicke** reported on a stroke patient that had a lesion in the left hemisphere, but with a different effect (in a different area), Receptive Aphasia, impairment of the comprehension, but can produce sentences (often wrong, but they're not aware of it).

Homunculus → the dimension of each part represents the amount of neurons in that area.

The study on aphasias led to the rise of molecular studies on the different parts of the brain to identify if the functional differences led to different structural features. Brodmann used tissue stain technique to identify different cells in neurons. Cytoarchitectonics: study of the cellular architecture → found 52 different regions. Brodmann categorization is still the atlas. Golgi invents “silver method” for staining neurons, using silver chromate to impregnate individual neurons. This helps to identify individual neurons. Using Golgi method, Cajal was the first to identify the unitary nature of the neuron → rise of the neuron doctrine: the nervous system is made up of individual cells/entities, discrete. Distinction lies in the number of processes arising from the cell body.

Principles of neural organization:

1. **Dynamic Polarization** → the electrical signal within a nerve cell flows from the receiving side of the cell to the triggered region, no other direction is possible
2. **Connectional Specificity** → there are specific points in the cell where neurons connect.

Fights on the nature of synaptic transmission. Fight between Electrical or Chemical. Chemical synapses can transport more information related to the chemicals they release, but they are slower than an Electrical Synapsis, therefore it depends on urgency.

The debate between the Aggregate Field Theory and Localizationism is still going strong today, though the core issue is that *the knowledge of the part cannot go without the knowledge of the whole.*

Cognitive Science Side

Also known as the Psychology side.

This starts with the Greeks, who asked themselves the questions: "Where are the most abstract features of personalities in our body?". There were two main philosophical schools:

- Cardio-centric (heart)
- Encephalo-centric (brain) → supported by Theon of Croton and Hippocrates

And then there was Rene Descartes and Dualism. The mind and the body are distinct and separable, and the body is a machine controlled by the soul and the soul is in the pineal gland, the only non-bilateral brain structure. This was opposed to monism, where Body and Mind are not separable, and the body produces the mind.

In 17th Century Rationalism (all the knowledge can be gained using reason alone, truth is intellectual and not sensory, they replaced religion) and Empiricism (knowledge is not something abstract, but comes from sensory experience - experience shapes who you are and what you believe in. In the beginning you start as a blank slate and then through direct sensory experiences you produce simple ideas and concepts. When these simple ideas interact and associate, complex ideas are created). School of Empiricism gives rise to one of the major schools of experimental psychology: Associationism and Behaviorism (Reinforcement Learning)

Associationism → individual experiences are associated to produce learning. Ebbinghaus: mental processes , like memory, can be measured and analyzed. Thorndike publishes the first statement about the nature of associations: Law of effect → a response followed by a reward will be repeated. If there's no reward, or even if there's punishment the response would disappear.

Behaviorism → Watson: psychology can be objective only if it were based on observable behavior (rejection of Ebbinghaus's theory). What's not observable is not measurable, and thus you can't do research with. Learning is the key, with everybody having the same neural equipment on which you can build learning and experience. (Little Albert Experiment →

instill fear about rats in a baby. Instills also a generalized aversive response on similar objects resembling the rat)

Cognitivism → all mental processes can be studied because the brain gives rise to them. Penfield (Montreal neurologist) leader of the school, invents a procedure to treat epilepsy based on surgically destroying the neurons responsible for the uncontrolled firing that led to seizures. He would use electrical stimulation to observe the behavior of the patient and determine which parts he could safely destroy and which preserve. Obtained maps of the sensory and motor cortices in the brain. Another researcher, Hebb, studied with Penfield the effects of brain surgery: learning has a biological basis: "Cells that fire together, wire together". Experiences shape the way neurons are connected, and synaptic communication becomes more efficient. It's about how the neurons are connected to each other. Concept of plasticity: the consequence is on the therapeutic approach to brain injury recovery, so we can train the brain to recover functions we have lost (to a certain extent, though). This theory is used in the design of artificial neural networks. Milner: Penfield's patients complained about memory loss after surgery → anatomical and physiological proof that there are multiple memory systems and the hippocampus, which worked for memory consolidation and short-term, and the extent of the deficit depended on how much of the medial temporal lobe was removed (the more posterior it was the more severe it was the amnesia and only bilateral resection of the hippocampus resulted in severe amnesia) - NB: the hemispheres are symmetrical. Procedural memories rely on basal ganglia and the cerebellum, while the hippocampus handles the declarative memory (as seen in the project HM).

Chomsky → Fall of behaviorism and associationism: you can't explain all learning through these two, citing language as something that cannot be explained through these theories. Innate and universal.

Back to Cognitive Neuroscience

To sum up: cognitive neuroscience studies how cognitive functions rise from neural system. Brain enables the mind, but the how? is what the cognitive neuroscience tries to answer to. Major goal is to understand the neural representations of mental processes and how ones rise from the others.

The Nervous System: Anatomy and Physiology

@March 1, 2023

The mantra now is "structure influences the function". Primary division of the Nervous System in Central Nervous System, composed by the Brain and Spinal Cord, and Peripheral Nervous System spreads from the spinal cord and innervates the whole body.

Individual Cells

- Nerve Cells or Neurons → transmit information across the system
- Glial Cells → support the activity of the neurons. An example is the oligodendrocyte that covers the axon of the neuron and the Schwann cell (same function, but in the PNS). This coverage is a physical protection for insulation, increase in transfer speed (electrical signal doesn't miss pace). Also protection from pathogens and external substances (microglial cells are part of the immune system). Astrocytes are for nourishment

Neuroglia

The term comes from "Nerve Glue" as they were originally thought to be supporting the neurons from a merely structural point of view. They actually support also nourishment, immune, signaling functions. They outnumber neurons 1:2 to 1:10.

- Microglia → immune system cells. They act as protectors: identify if something's gone wrong and initiate a response that removes the toxic agent and/or clears away the dead cells (phagocytose them). In neurodegenerative disorders (e.g. Alzheimer's), microglia can become hyperactivated and promotes neuroinflammation that leads to the formation of toxic protein deposits (amyloid plaques outside the neurons and neurofibrillary tangles inside the neuron). On a macro-level, you see a reduction of volume, and the indentations are much bigger, due to cell death (missing tissue).
- Astrocytes → important for nourishment. Star-shaped and symmetrically round. They surround neurons and connect them to the blood vessels. They also regulate the presence of ions and neurotransmitters in the extracellular space. Astrocytes also support synaptic communication (recent research topic) and maintain the blood-brain barrier between the tissues of the CNS and the blood (protection from infections/external substances - there are two other layers: skull and meninges).
- Signaling support → Oligodendrocytes (CNS) and Schwann cells (PNS). They provide the insulating layer around the neurons, whose membrane cells are mainly made of liquids (poor conductors, famously). They produce thin sheets of myelin that are wrapped concentrically around the axons of neurons for rapid conduction. The myelin is white (that's why the term "white matter" has been coined). A difference is that one oligodendrocyte can myelinate more than one neuron while Schwann cell myelinates only one axon.
 - if myelin is damaged, the signal has problem in maintaining speed/amplitude and this might lead to issues (more or less impairing according to the disease) → Multiple Sclerosis: the symptoms are from damage to oligodendrocytes.

Autoimmune disease: the immune system attacks (mistakenly) healthy parts of the body.

Neurons

Huge variety and more than huge presence. Each neuron receives and gives rise to thousands of connections (this is the key). These connections are shaped from experience. Some of these connections are formed nearly a meter away from the cell body of the neuron. They vary in form, location and interconnectivity.

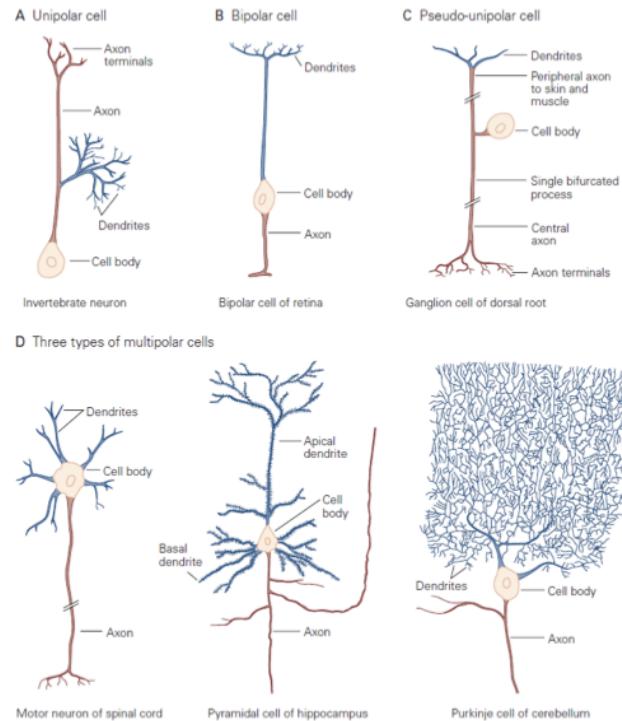
Functions:

1. Take information in
2. Decide what to do with the information that's in
3. At the end of the axon, pass it along to the other neurons

The transmission is by changes in electrical levels of activity.

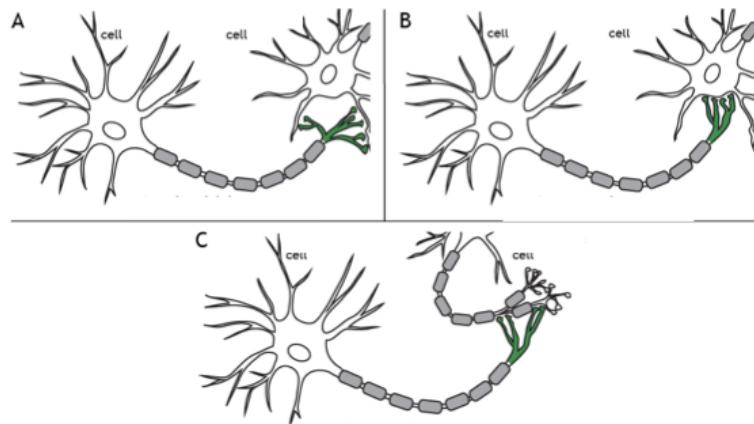
Structure:

- Components common to eukaryotic cells
 - Cell membrane → membrane that separates the intracellular space from the extracellular space
 - Cytoplasm → intracellular fluid, made up of a combination of ions (play a role in electrical transfer - K^+ , Na^+ , Cl^- , Ca^{2+}) and different other molecules such as proteins.
 - Extracellular fluid → bath where the neurons sit, made up of a mixture of the same type of ions found in the cytoplasm
 - Cell body or soma → it contains the nucleus (which is present in all but red blood cells). Metabolic center of the cell. It has many structures inside: nucleus (DNA) and other structures important for the functioning of the cells (e.g. endoplasmic reticulum, where proteins are synthesized).
- Components unique to neurons, which have unique functions in generating signals and communicating with other nerve cells.



- **Dendrites** (many) → main apparatus for receiving incoming signals from other nerve cells. They can take many different forms, according to the type and location of the neuron. They are the receiving end
- **Axon** (only one) → The output side of the neuron. It doesn't only communicate with neurons, but also other cell types, i.e. muscles. Can convey electrical signals over varying distances in length.
- **Synapses** → specialized structures at the end of the axon. Pass the signal between neurons/neurons and cells, it also includes sections from the dendrites of the receiving end (postsynaptic cell). They pass chemical and electrical signals, enabling communication between neurons.
 - Presynaptic Cell → nerve cell transmitting a signal. It's composed of presynaptic terminals or nerve terminals (i.e. specialized enlarged regions of its axon's branches)
 - Postsynaptic Cell → cell receiving the signal
 - Synaptic Cleft → narrow space separating the pre-synaptic and post-synaptic cells

I can have different types of synapses:



- Axosomatic → synapse made onto the soma or cell body of a neuron (B)
- Axodendritic → synapse made onto the dendrite of another neuron (A). It is the most common type.
- Axoaxonic → synapse made by one neuron onto the synapse of another neuron. They mediate presynaptic inhibition and facilitation (C)

I can classify neurons also based on their major functional category:

- Sensory → carry information from the peripheral sensors to the CNS for the purpose of both perception and motor coordination. Five main senses, same signals.
- Motor → carry commands from the CNS to the PNS to generate motor output.
- Interneurons → Neurons that mediate impulses between sensory and motor neurons

https://youtu.be/qPix_X-9t7E

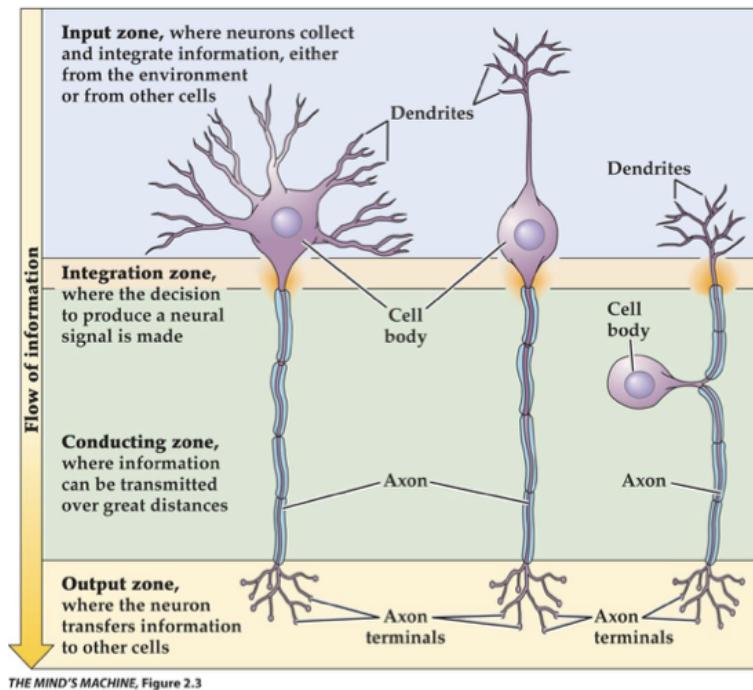
Information Transfer within a Single Neuron

Information can be transmitted:

- Within a neuron → received at synapses on dendrites, we have an input region, an integration region that integrate the information received (it is worth being passed along/ignored etc). There's a region that conduct the information (axon) and a transmission region (output) - synapses. Four types of signal:
 1. Input signal (receptive)
 2. Trigger signal (summing or integrative)

3. Conducting signal (signaling)

4. Output signal



The Signaling is organized in the same way in all nerve cells

- Between a neuron and another cell (that can be also non-neuronal)

Signaling within a neuron

The transferring of information involves transient changes in the electrical state of the neuron and it is produced by temporary changes in the electric current into and out of the cell. It is all possible because of the resting membrane potential (voltage imbalance between the inside and outside of the cell: the inside is $\approx 70\text{mV}$ more negative). This means that the neuron has at its disposal a battery of sorts, whose stored energy can be used to do work.

The imbalance arises from the asymmetric distribution of ions across the neuron's cell membrane. As the cell desires equilibrium of concentrations, we have two mechanism (one active and one passive):

- Sodium-Potassium pump (Na^+/K^+ pump) → needs energy to work (works against the concentration gradient)

Transports out more sodium and inside more potassium, creating a concentration gradient (there's more sodium outside than inside - more potassium inside than outside)

- **Ion channels** → they let the ions pass passively (only in the direction where there's less similar charge to them, along with the concentration gradient)

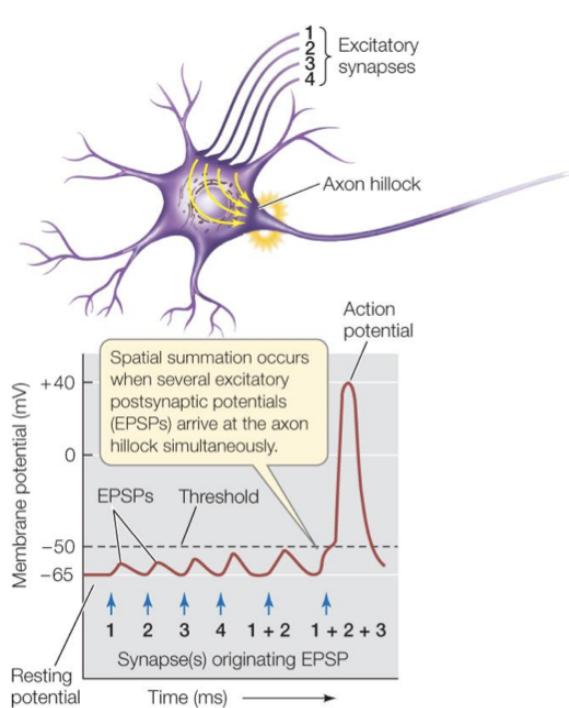
They don't need energy and let specific ions pass according to concentration gradient (they're the mechanism to restore balance in the cell - though there are more K channels than Na channels). The pump fights against it.

There's both a concentration gradient that push the potassium out and an electrical gradient that attracts the potassium in until they reach electrochemical equilibrium ($-70mV$).

I have different electrochemical forces that cause the inside of the cell to have a more negative potential than the outside. The imbalance is the energy source used for signaling. The resting membrane potential can be quickly and significantly altered, with the alteration acting as signaling mechanism.

Postsynaptic potential can alter the imbalance:

- Depolarization (decrease) → excitatory postsynaptic potential: enhances the ability to generate an action potential
- Hyperpolarization (increase) → inhibitory postsynaptic potential: reduces the ability to generate an action potential



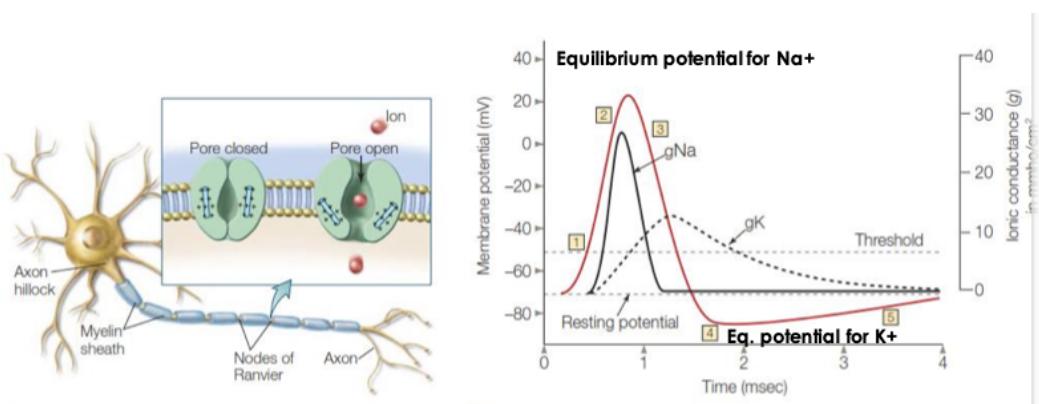
An individual neuron sums excitatory and inhibitory postsynaptic potentials received at spatially separate synapses. When the sum of the potentials present at one time depolarizes the axon hillock to threshold, the neuron generates an action potential.

The Action potential is the binary part, the signal summation is the analogical.

Need a system to ensure that information is passed efficiently, fast and far. The maximum distance a passive current flows is 1mm, with the longest axons going for meters.

The action potential is actively propagated along the axon.

The action potential is actively propagated along the axon, and it has a self-regenerative feature, its amplitude does not diminish by the time it reaches the axon terminal.



Neurons evolved a mechanism to regenerate and pass along the signal initiated by the PSPs. It is an active membrane mechanism known as Action Potential. It is a rapid depolarization and re-polarization of a small region of the membrane caused by the opening and closing of the ion channels.

When the depolarization caused by the summation of the PSPs is strong enough (hits the threshold of -55mV) the action potential is triggered. The depolarization causes the voltage-gated sodium channels to open, with sodium flowing rapidly inside the cell and neutralizing the negative charge inside. This goes on until the Equilibrium Potential for Na^+ is reached (40mV). At this point, the voltage-gated potassium channels open to let potassium flow out of the neuron to balance the concentration gradient. This process stays until the membrane potential reaches the Equilibrium Potential for K^+ , which is actually more negative than the resting state (hyper-polarization). At this point the K^+ channels close.

During hyper-polarization the voltage-gated Na^+ channels cannot open (refractory period). Functional consequences:

1. The cell has a limit in the number of action potential that can generate in a given time
2. Unidirectional current flow from the axon hillock to the axon terminal → the current can't reopen the channels that generated it, so it can only flow ahead.

Confirms the Principle of Dynamic Polarization made by Cajal.

<https://youtu.be/LcO9YU-Pdws>

The action potential needs to travel far and fast. This is ensured by Oligodendrocytes and Schwann Cells, which produce thin sheets of myelin that wrap around the axon of neurons. They provide the insulation material (myelin is made of lipids, notoriously bad conductors), which increases resistance to voltage loss, allows rapid conduction of electrical signals and

reduces the need for frequent regeneration.

In myelinated axons, action potentials occur only at the nodes of Ranvier, zones where myelination is interrupted and the channels and pumps are actually located. The appearance is of a saltatory conduction.

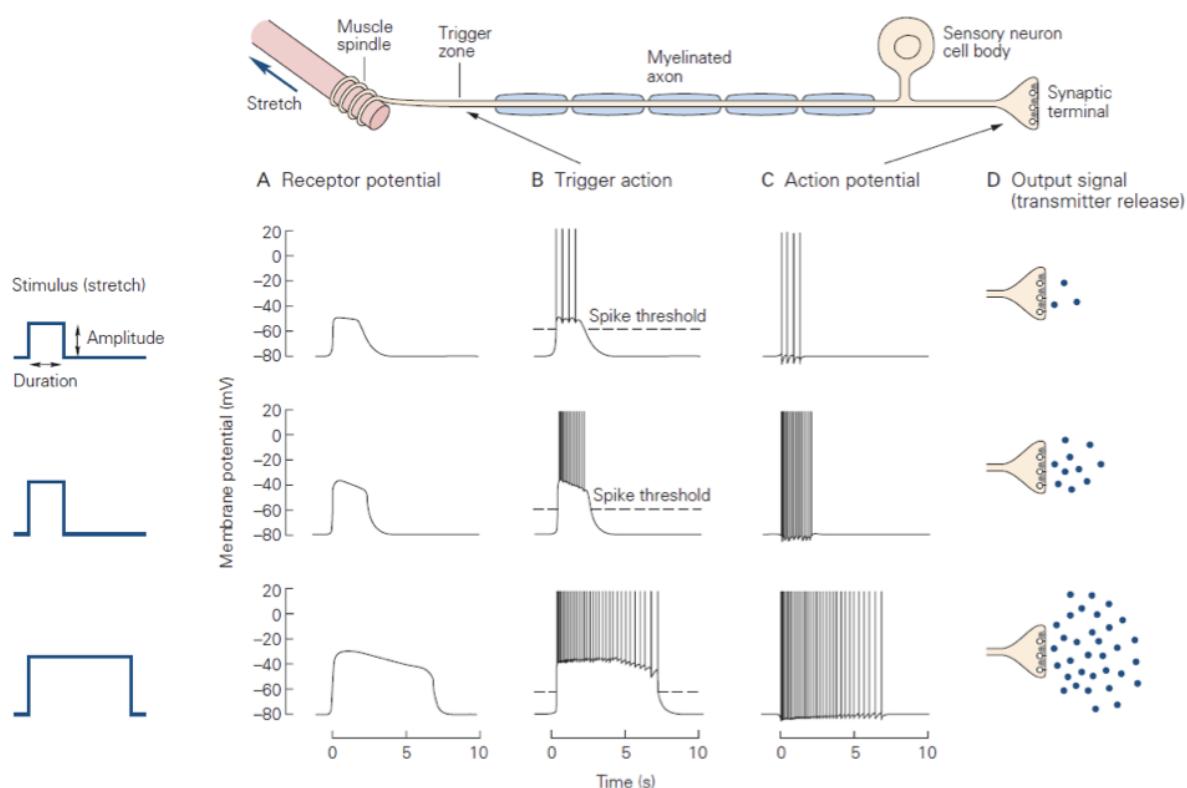
@March 8, 2023

Properties of APs:

1. Threshold for initiation → only “meaningful” information lead to the Action Potential.
2. Conducted without decrement → amplitude of AP is constant throughout the axon, as AP is regenerated at each node of Ranvier (self-regenerating properties)
3. Refractory period → Signal can only travel through the channel but can’t go backwards.
Neurons don’t get overloaded. This limits the frequency at which a nerve can fire an action potential.
4. All-or-none nature → the AP is a binary signal; if the threshold is passed, that’s it.

The thing we can modulate to determine the stimulus’ strength is the frequency of the signal.

Regardless of the structure, there are 4 regions that encode 4 types of signals.



Each of the neuron's four signaling regions produces a characteristic signal

- a. At the input region, the input signal is graded in amplitude and duration proportional to the amplitude and duration of the stimulus.
- b. The trigger zone sums the PSPs and establish whether to generate the AP. The action potential is generated only if the input signal exceeds the voltage threshold for initiation. Once the threshold is surpassed, the AP is generated and any further increase in amplitude can only increase the frequency of APs, while the duration of the input determines the duration of the train of action potentials. The graded amplitude and duration of the receptor potential is translated into a frequency code in the APs generated at the trigger zone. All APs produced are propagated along the axon.
- c. The conductive region transmits action potentials. They are all-or-none and the frequency and duration of firing represents the information carried by the signal.
- d. Output region produces the output signal responsible for synaptic communication. At chemical synapses, the frequency of action potentials determines exactly how much neurotransmitter is released by the cell. At electrical synapses, the signal is directly transmitted.

https://youtu.be/OZG8M_ldA1M

From Individual Neurons to Neural Circuits

What happens when a combination of neurons have to communicate between each other.

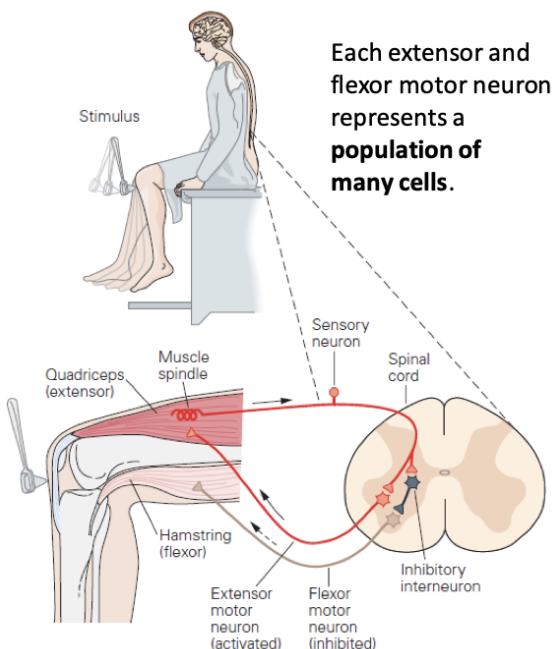
We call Neural circuits groups of interconnected neurons that process specific types of information.

1. Neurons that take in informations (sensory neurons - input)
2. Evaluation of the input at synapses or within one or a group of neurons (process, integration)
3. Output/reaction (Motor Neurons - output).

Every behavior is mediated by specific sets of interconnected neurons, and every neuron's behavioral function is determined by its connections with other neurons.

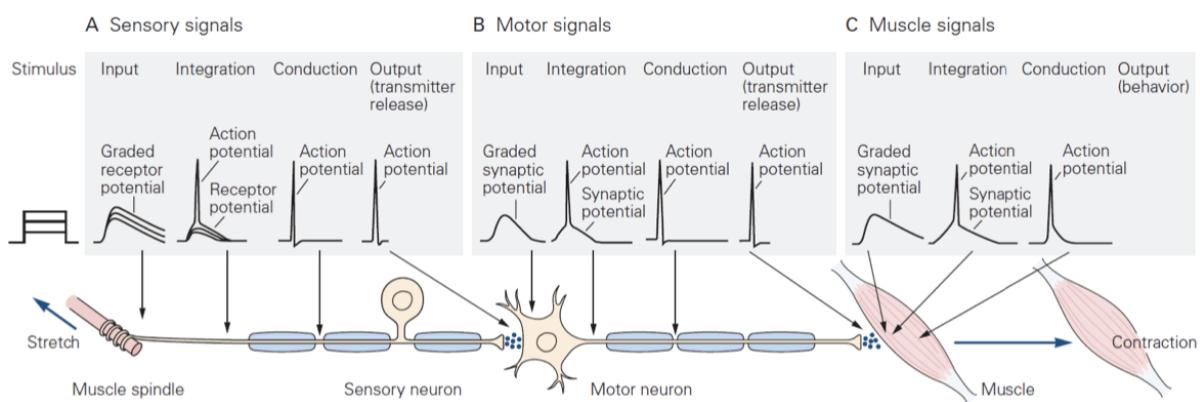
The Knee-Jerk Reflex

I have two motor neurons (Flexor and Extensor) because muscles usually work in pairs with opposing effects (agonist and



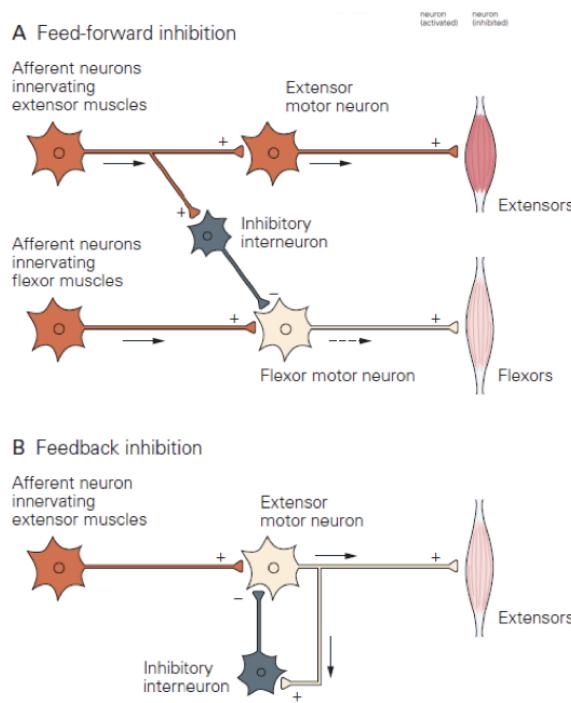
antagonist). The sensory neuron synapses both on the Extensor Motor Neuron firing it up and on the inhibitory interneuron that inhibits the Flexor Motor neuron. This allows the reflex to be produced.

Sensory information is conveyed to the CNS (spinal cord) from the muscles. Motor commands from the CNS are issued to the muscles that carry out the knee jerk; at the same time inhibitory commands are issued to motor neurons that innervate the antagonist muscles, providing coordination.



- Divergence and Convergence
 - Divergence (input stages): one neuron activates many cells → one sensory neuron can have a widespread effect
 - Convergence (output stages): many neurons activate a single cell → ensure you need enough information in order to produce an output
- Excitatory and Inhibitory Neurons
 - Excitatory Neurons: output increases probability the post-synaptic neurons will fire it. Signal depolarizes the membrane and reduce the threshold for AP
 - Inhibitory Neurons: opposite effect.
 - Feed-Forward Inhibition → It is useful as it enhances the effect of the active pathways by suppressing the activity of pathways mediating opposing actions.

- Feed-Back Inhibition → the inhibitory signal inhibits the same neuron that activates it. It is useful as it dampens the activity within the stimulated pathway and prevent it from exceeding a certain critical level



No complex human behavior is initiated by a single neuron. You need at least a circuit. The control of behavior has three components: a sensory input, an intermediate processing, a motor output. In vertebrates each component is mediated by a single group or several distinct groups of neurons and has multiple neural pathways that simultaneously provide the same or similar information.

This form of Parallel Processing increases speed and reliability of functions within the central nervous system. It's kind of a fail-safe.

Plasticity

Chemical synapses are functionally and anatomically modified through experience and learning throughout life. These changes can be long-term (structural → rise to further physiological changes that lead to anatomical alterations) or short-term (functional → increase or decrease the effectiveness of existing connections). Principle of “use it or lose it” = synapses you don’t use anymore are lost

https://www.youtube.com/watch?v=w6AfzCNDmbY&ab_channel=StarStuff

Watch only the Phantom Limb Pain section 😊

Information Transfer between two neurons

(a.k.a. the Synaptic Process)

Within a neuron → transferring information involves changes in the electrical state of the neuron → neuronal spikes

Between neurons → information occurs at synapses, both of electrical and chemical nature.

The structure is different. In Electrical Synapses neuronal membranes are almost in direct contact (pores connect cytoplasms) in a specific area called gap junction.

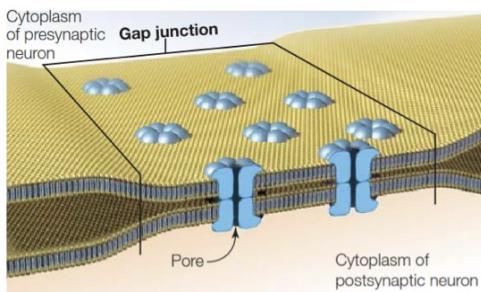


FIGURE 2.14 Electrical synapse between two neurons.

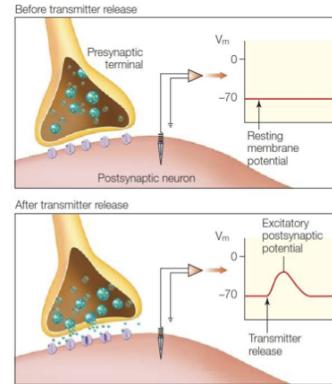


FIGURE 2.13 Neurotransmitter leading to postsynaptic potential.

In Chemical Synapses there's no structural continuity (we have synaptic cleft). This implies that electrical synapses are faster.

- Electrical Synapses

The gap junction channels create pores connecting the cytoplasm of the two neurons, which are iso-potential. Electrical changes in one neuron are reflected instantaneously in the other.

Less specific, can't modulate (what arrives, passes) and less plastic. It allows for synchronous working of different groups of neurons. Can be bi-directional by structural definition (in theory).

- Chemical Synapses

Based on diffusion of neurotransmitters across the synaptic cleft. We have:

1. Neurotransmitters: chemical substances that bind receptors in the postsynaptic membrane of the target cells

2. Presynaptic terminals: specialized swellings of the axon, which typically contain vesicles

3. Synaptic vesicles: vesicles filled with several thousand molecules of neurotransmitter.

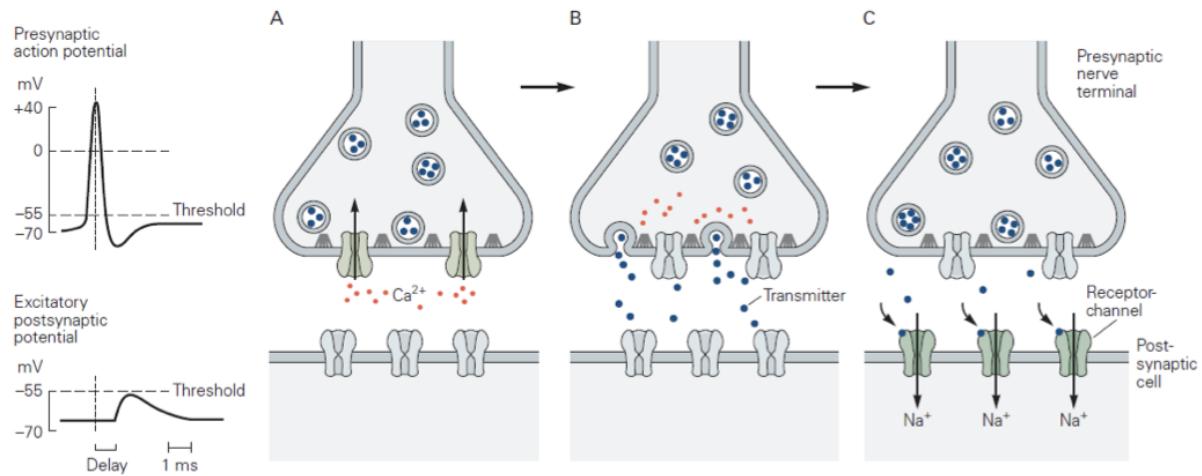


Figure 8–8 Synaptic transmission at chemical synapses involves several steps. The complex process of chemical synaptic transmission accounts for the delay between an action potential in the presynaptic cell and the synaptic potential in the postsynaptic cell compared with the virtually instantaneous transmission of signals at electrical synapses (see Figure 8–2B). A. An action potential arriving at the terminal of a presynaptic axon causes voltage-gated Ca^{2+} channels at the active zone to open. The gray filaments represent the docking and release sites of the active zone.

B. The Ca^{2+} channel opening produces a high concentration of intracellular Ca^{2+} near the active zone, causing vesicles containing neurotransmitter to fuse with the presynaptic cell membrane and release their contents into the synaptic cleft (a process termed *exocytosis*).

C. The released neurotransmitter molecules then diffuse across the synaptic cleft and bind specific receptors on the postsynaptic membrane. These receptors cause ion channels to open (or close), thereby changing the membrane conductance and membrane potential of the postsynaptic cell.

Neurotransmitters can be inactivated through active re-uptake, enzymatic breakdown or diffusion away. They don't stay in the receptors forever.

Many different kind → different neurotransmitters can open different channels and have different effects. They can have both excitatory or inhibiting effect (although usually they have a typical effect) depending on where's hitting.

Slower, but plastic, possible modulation, specific.

<https://youtu.be/VitFvNvRIIY>

From Neural Circuits to Neural Systems

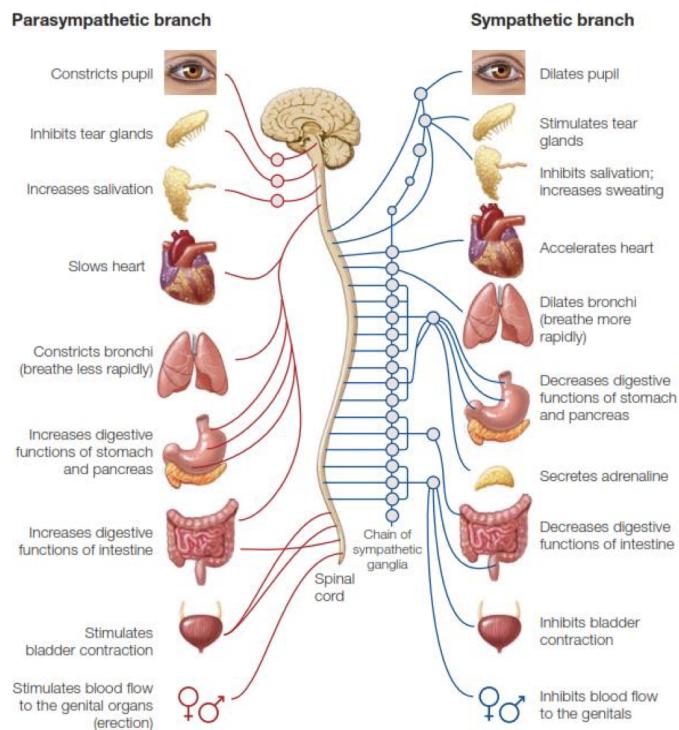
Neural Systems = combination of neural circuits

Peripheral Nervous System

Nerves → bundles of axons and glia. Ganglia → clumps of nerve cell bodies outside the CNS. Transmits sensory information to the CNS and carry motor output from the CNS.

Supplies to the CNS information about external and internal environments. It can be subdivided:

- Somatic Nervous System → sensory neurons that receive information from the skin, muscles and joints. Receptors provide information about proprioception, touch, pressure at the body surface. They transduce different types of physical energy into electrical signals.
- Autonomic Nervous System → Mediates visceral sensation, motor control of viscera, vascular system and exocrine glands
 - Sympathetic → fight-or-flight response by stimulating adrenal glands to release adrenaline. Uses norepinephrine
 - Parasympathetic → maintain homeostatic equilibrium. Uses acetylcholine. It is antagonist to the Sympathetic NS.



- Enteric → control smooth muscle of the gut

<https://youtu.be/71pCilo8k4M>

Central Nervous System

Made of brain and spinal cord. As it is very important, it is well protected. The most outer layer is the skull and vertebrae.

Three protective membranes:

- Dura Mater → the most outer, thickest
- Arachnoid Mater
- Pia Mater → inner membrane, the most delicate, firmly adhere to the brain surface tissue

In addition, liquid protection: Cerebrospinal fluid between the Arachnoid and Pia Mater. It runs around all the CNS (also in the brain ventricles, cisterns, sulci, central canal of the spinal cord). It allows the brain to float (if you have mechanical damage to the brain, reduces the shock and offset the pressure).

Another barrier against pathogens and toxins is the blood-brain barrier, situated between the brain's blood vessels and the components of brain tissue. Sometimes can cause problems (it might be necessary to inhibit it to assume particular brain drugs).

- **Spinal Cord**

It is a relay station, takes in sensory information from the body's peripheral sensory receptors and conducts the final motor signals from the brain to muscles. It is enclosed in the vertebral column, a stack of separate bones (vertebrae) that extends from the base of the skull to the coccyx. It is divided in 31 segments, with each segment having a right and left spinal nerve. Sensory axons through the dorsal root and Motor axons through the ventral root.

Pathways in the central nervous systems are bilaterally symmetrical and cross over to the opposite side of the brain and spinal cord. Left hemisphere controls the right side of the body and Right controls left side.

Neurons are bunched together in Nuclei (compact arrangement of nerve cell bodies and their connections, located in the entire CNS) and Layers (thin sheets, folded across the surfaces of the cerebral hemispheres, found in the cerebral cortex)

- **Brain**

Composed of 6 Subdivision and it is mostly symmetrical along the midline: each subdivision is found in both hemispheres with slight bilateral differences

- **Medulla, Pons, Midbrain (Brain Stem)**

Regulates basic life functions (the most primitive - blood pressure, respiration, sleep). Injury is life threatening.

- **Cerebellum**

Responsible for movement and stability. Important also for relaying information with the cerebral cortex, language and cognitive functions (more recent research). It contains far more neurons than any other subdivision of the brain

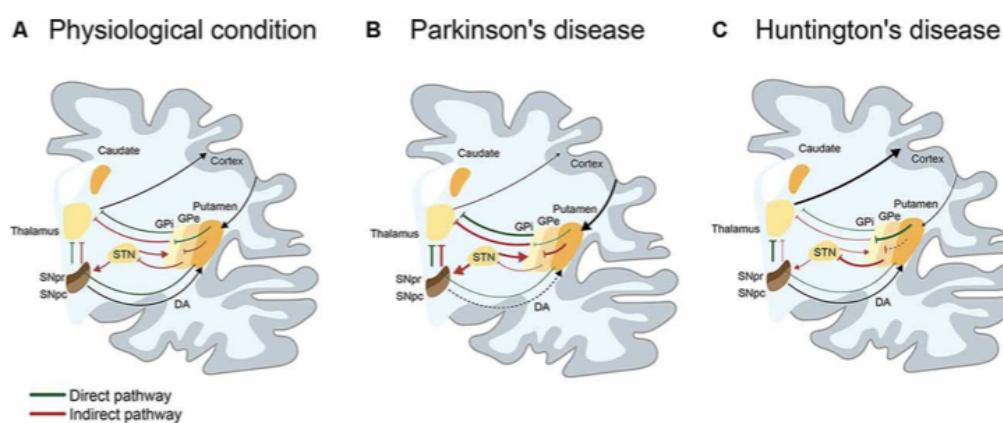
- Diencephalon

- Thalamus → gateway to the cortex. Essential link in the pathway of sensory information from the periphery to sensory regions of the cerebral hemispheres. Determines which sensory information reaches the neocortex.
- Hypothalamus → ventral to the thalamus, link between the nervous and endocrine system. Regulates body temperature, thirst, hunger, circadian rhythm. Important for motivation and maintaining and regulating behaviors related to aversive or rewarding stimuli

- Cerebral hemispheres/Telencephalon

Largest part of the human brain. Consists of:

- Basal Ganglia → collection of subcortical nuclei, receive inputs from sensory and motor areas and send output largely through the thalamus to the frontal lobe. They have a crucial role in motor control. Parkinson's and Huntington's Disease.



Physiologically, the direct pathway participates in the activation of movement, while the indirect pathway participates in inhibition of movement. In Parkinson's Disease, the loss of dopaminergic neurons induces an overactivation of the indirect pathway (hypokinesia). In Huntington's Disease we have the opposite effect, with uncontrolled movements called Chores.

Also crucial role in Reinforcement Learning because rich in dopamine receptors.

- Amygdala → Small, almond shaped structures in the medial temporal lobe adjacent to the anterior portion of the hippocampus. It is a collection of 13 nuclei. It is involved in attention, perception, value representation, decision making, learning and memory
- Hippocampus → small, curved formation. Crucial for memory formation and spatial memory

@March 15, 2023

- Cerebral Cortex → made mostly of layered neurons/grey matter (neuronal cell bodies). Nuclei in subcortical areas underneath the cortex. We have two symmetrical hemispheres connected by the corpus callosum, where axons travel to connect the hemispheres. The total surface is up to 2.4 squared meters. The folded structure in sulci (crevices) and gyri (crowns) brings neurons to a closer three-dimensional relationship to one another and it increases connections as axons don't necessary have to follow the infoldings, but can use shorter paths through the white matter (reduces axonal distance and neuronal conduction time).

Brodmann identified 52 distinct regions, a division still used in cytoarchitectonics, that reflects also functional differences. In the cortex the neurons are layered; the neocortex, in red, includes areas like the primary sensory and motor cortex and association cortex. In the neocortex, we can identify six different layers (sheets of neurons neatly stacked on top of each other), each with differently-shaped neurons (that reflects the slightly different functioning). Deeper layers mature earlier than more superficial ones, which are thought to be involved in higher cognitive functions. We can identify also vertical columns in the neocortex; the neurons in a column synapse with neurons both above and below them → functional units.

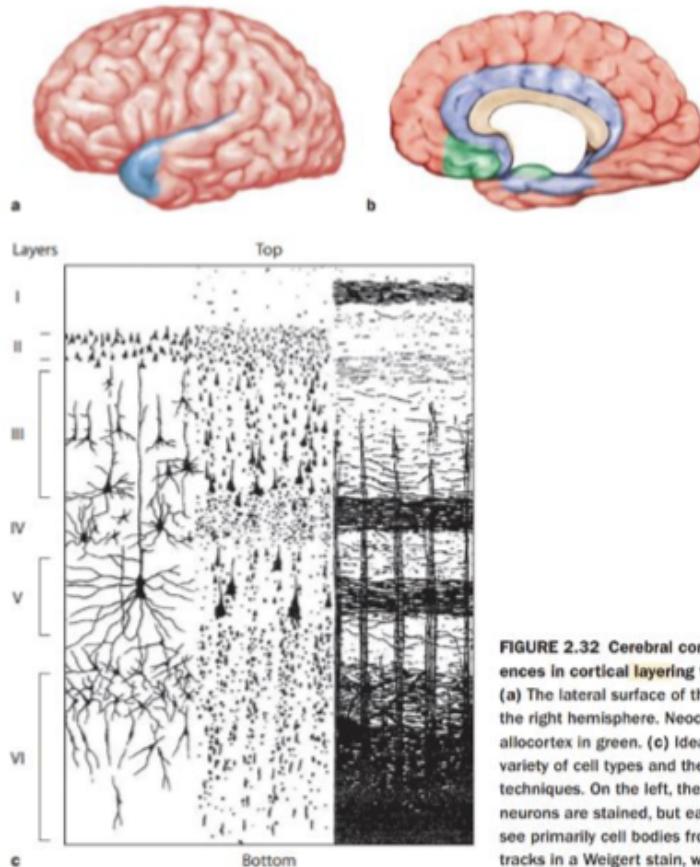


FIGURE 2.32 Cerebral cortex, color-coded to show the regional differences in cortical layering that specify different types of cortex.
(a) The lateral surface of the left hemisphere. **(b)** The medial surface of the right hemisphere. Neocortex is shown in red, mesocortex in blue, an allocortex in green. **(c)** Idealized cross section of neocortex showing a variety of cell types and the patterns of three different types of staining techniques. On the left, the Golgi preparation is apparent: Only a few neurons are stained, but each is completely visualized. In the middle, we see primarily cell bodies from the Nissl stain. On the right, we see the fiber tracks in a Weigert stain, which selectively stains myelin.

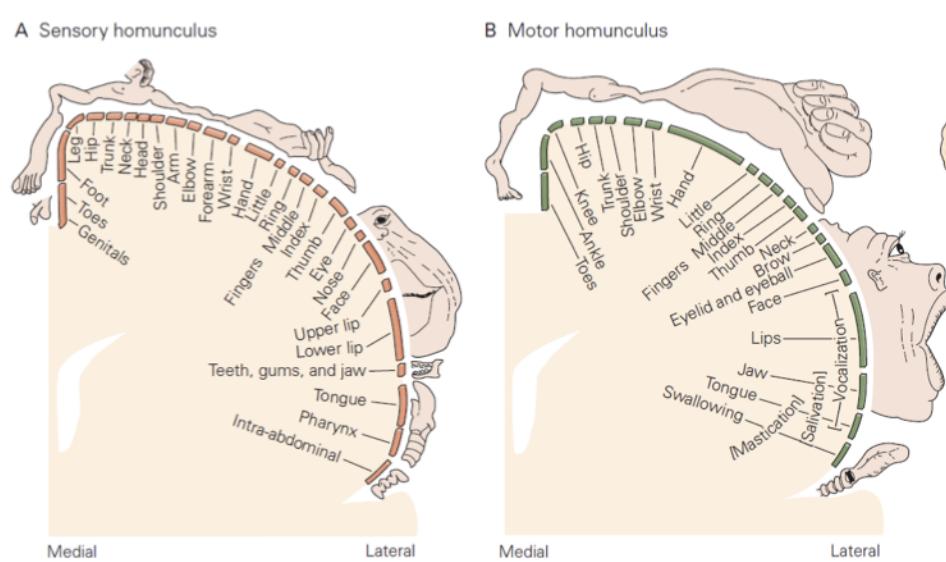
At a macro-level, the cortex is divided in 4 different lobes. Different lobes have different functions. Within each lobe, the functional system is hierarchically organized (especially in Occipital Lobe - Visual Processing).

Functional Division:

- Frontal Lobe → it is the one that matures last (takes longer to mature also during development). Section of the neocortex which is ventral to the central sulcus.
 - Motor Cortex → neurons in the motor cortex descend to the spinal cord and activate muscles. Involved in the planning and execution of movements.
 - Prefrontal Cortex → important for higher cognitive functions (planning, inhibitory control, decision making, motivation...). It is the part which is most developed compared to primates
- Parietal Lobe → between the central sulcus and the parieto-occipital sulcus. Receives sensory information from the outside world, from the inside (within the body), memory. It integrates all these information.

- Somatosensory Cortex → ventral to the motor cortex, divided by the central sulcus. Information about touch, pain, temperature, proprioception. Uses receptors from the peripheral to the central NS (somatosensory relays of the thalamus).
- Higher Order Sensory Areas → multi-modal integrations.

Somatosensory cortex and motor cortex can be organized to have a representation (I can draw a neural map).



Homunculus. The proportion/extend to which each body part is represented reflects the density of the innervation of the body part

- Occipital Lobe →
 - Visual Cortex → processes visual information. The primary visual cortex begins the cortical coding of visual features (luminance, spatial frequency, orientation, motion). We can draw retinotopic maps (orderly mapping between spatial location and neural representation).
- Temporal Lobe → Includes the auditory cortex, which do sound processing (from the cochlea through the subcortical relays to the thalamus to reach primary auditory cortex). It also has a map organization (tonotopic map => layout of the neurons based on sound frequency of different auditory information).
- Association Cortex → Information from multiple sources has to be integrated. Portion that is neither sensory nor motor. Each sense has an association area that integrate all separate features of a sensory input. It is responsible for all the high-end human abilities.

- White Matter → (axons and glial cells)

https://youtu.be/q8NtmDrb_qo

Learning, Decision Making and Reinforcement Learning

We make decisions each day, each moment. Optimal decision making consists of maximizing the reward while minimizing the losses/punishments. It is hard, as the actions can be divided in subproblems, and therefore the total long-term outcome depends on a series of actions, each with their local outputs. Also, the outcome might be delayed.

Reinforcement Learning Algorithms work to solve the Credit Assignment Problem: how do you distribute credit for success (or blame for failure) of a decision among the many component structures that could have been involved in producing it.

Animal Learning and Reinforcement Learning

Learning is defined as enduring change in response or behavior that occurs as a result of experience. It is long-term. Two broad classifications:

Non-Associative Learning

You learn about the properties of a single stimulus. After the repeated exposure of the participant to a single type of stimulus, one can visualize:

1. Habituation → the original response decreases. Over time, it becomes less likely to respond to the stimulus (e.g. fireworks).
2. Sensitization → Over time, more likely for the subject to respond to noxious stimuli (e.g. if you scratch yourself, at the beginning it might be pleasurable, but as the behavior goes on it becomes painful).

Associative Learning

The change in response and behavior is caused by the association of at least two stimuli or events: a neutral stimulus (the sound of reboot for Dwight or Penny's actions) and a reinforcer - innately meaningful (the piece of candy for Dwight and the chocolate for Penny). It doesn't depend on what the person that is learning does. Associative learning is the way in

which an animal learns to predict events and control them, in order to maximize rewarding outcomes and minimize aversive outcomes to increase chances of survival.

Reinforcers can be:

1. Primary (food, pain, social approval) → stimuli biologically prepared to elicit a response, they're biologically relevant and can be positive or negative.
2. Secondary (money, social media approval - mimics social approval, grades) → stimuli that come to elicit a response following associative learning; they become relevant following the associative learning. Can be positive or negative.

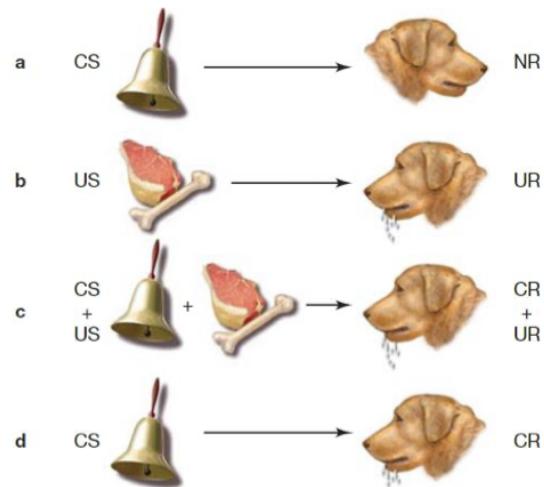
Associative learning is constrained by the biology of the organism. Animals generally learn to associate stimuli that are relevant to their survival. Conditioned taste aversion: occurs when certain tastes are associated with nausea/sickness (e.g. the bottle of Tequila). It develops poorly if a taste is followed by a stimulus that does not produce nausea and does not develop for visual or auditory stimuli that have been paired with nausea.

Two classes of Associative Learning:

1. Relationship between a stimulus and an outcome: Pavlovian Conditioning → Prediction Learning (I learn to predict events in the world)
2. Relationship between a behavior and an outcome: Instrumental/Operant Conditioning → Control Learning (learn how to control the world in order to get a specific outcome)

Pavlovian or Classical Conditioning (Prediction Learning)

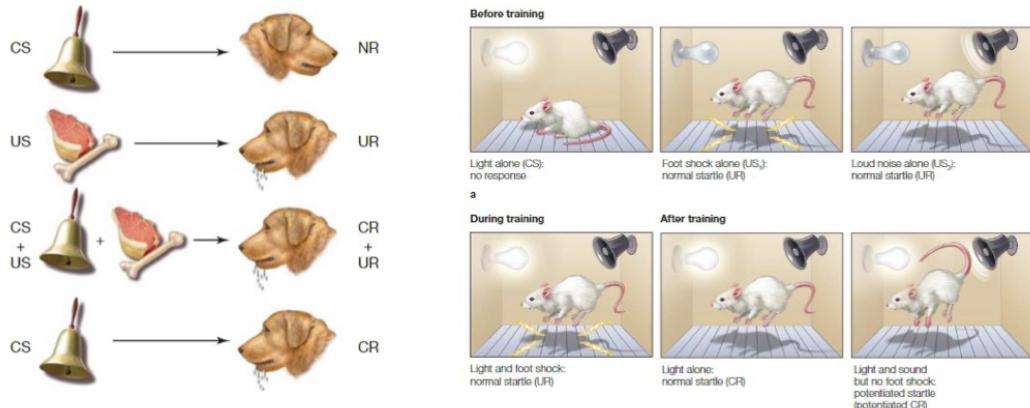
When a stimulus is presented that has no meaning to an animal, such as the sound of a bell (Conditioned Stimulus), there is No Response. In contrast, presentation of a meaningful stimulus like food (Unconditioned Stimulus) generates an Unconditional Response. When the sound is paired with the food, however, the animal learns the association; and later the newly Conditioned Stimulus alone can elicit the response, which is now called a Conditioned Response



Pavlov was actually studying the digestive system of dogs, not learning specifically.

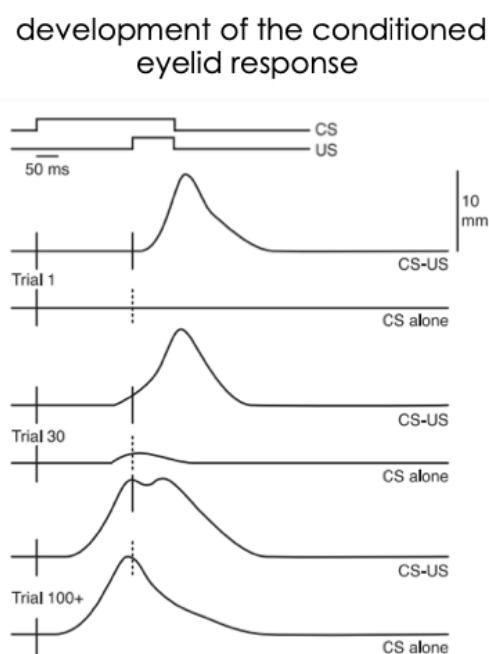
You pair the unconditioned meaningful stimulus with a neutral stimulus that does not have originally a meaning.

What he observed was that the response that was originally elicited by showing the food, was in the end elicited just by the bell sound that came before the food was shown.



Stimuli can be appetitive or aversive.

CR must be learned, UR takes place without any learning necessary. CR are anticipatory, predictive responses. They anticipate the CS. These responses can be behavioral, psychological, change in subjective experience.



Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.

The animal responds to the CS with a CR that prepares the animal for, or protects it from, the predicted US.

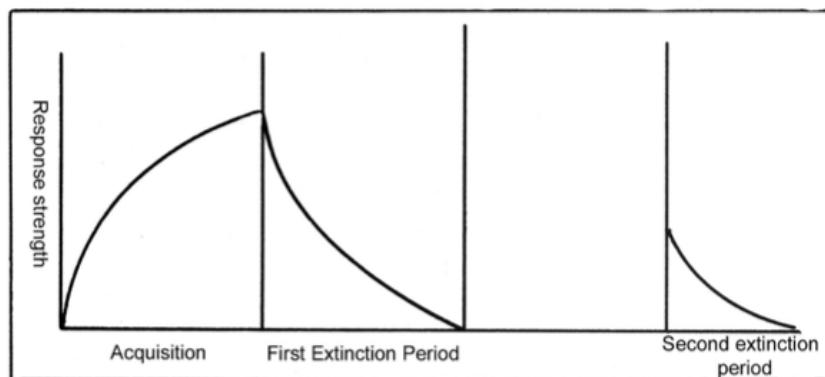
The tone comes to trigger a CR consisting of the nictitating membrane closure that begins before the air puff and eventually becomes timed so that peak closure occurs just when the air puff is likely to occur. This CR, being initiated in anticipation of the air puff and appropriately timed, offers better protection than simply initiating closure as a reaction to the irritating US.

Timing matters: we can have

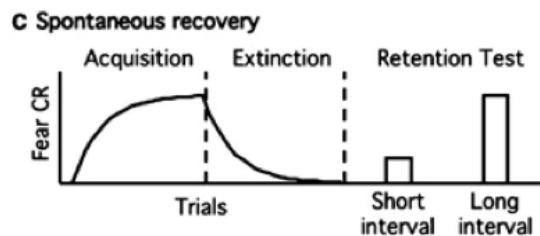
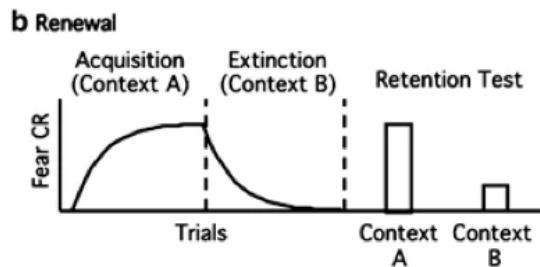
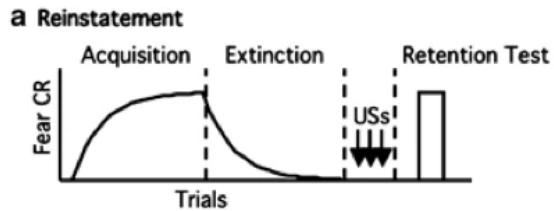
- Forward conditioning → the CS anticipates the outcome/reinforcer and gets to become an anticipatory stimulus for the US. Learning is most effective
- Simultaneous conditioning
- Backward conditioning

Also frequency matters:

- Continuous Reinforcement → the CS is reinforced with the US every single time it occurs, it is most effective when trying to teach a new association
- Partial Reinforcement → the CS is reinforced only part of the time. Associations are acquired more slowly with partial reinforcement, but the response is more resistance to extinction

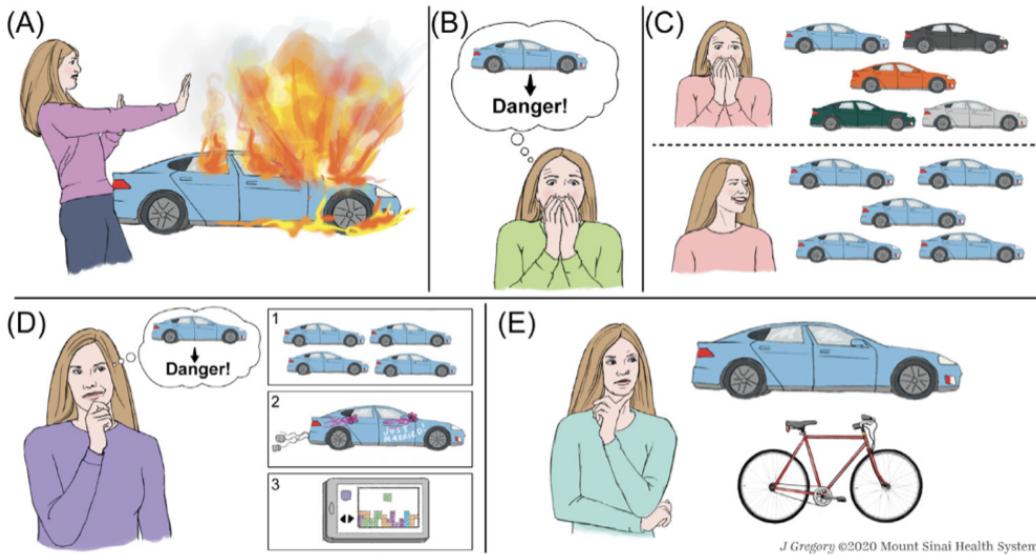


The probability of occurrence of a conditioned response increases if the CS is repeatedly presented with the US (**Acquisition**). Adaptive mechanisms ensure that an animal responds to cues that are meaningful to survival. The probability of occurrence of a conditioned response decreases if the CS is repeatedly presented without the US (**Extinction**). Adaptive mechanism: it ensures that an animal stops responding to cues that are no longer meaningful to it. After extinction, the original CR can return under specific circumstances.



Fear/Threat/Aversive Conditioning

Important process for survival. Fear Conditioning is a pervasive form of learning, as it can influence the subject's life (can also be maladaptive).

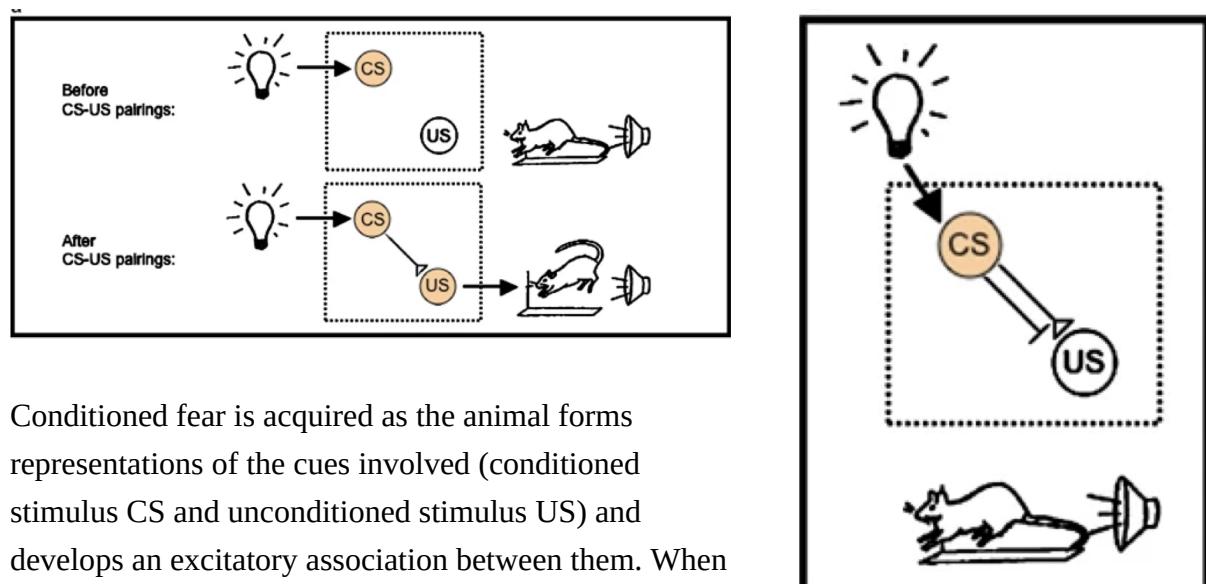


J Gregory ©2020 Mount Sinai Health System

Trends in Cognitive Sciences

Figure 1. The Stages of Threat Experience. Experiencing a life-threatening event, in this case an aversive emotional memory of a car explosion (A), may result in associative learning (B) where a neutral stimulus (the blue car) becomes threatening as it predicts danger (explosion). The learned association then competes with or influences new associations (C). For example, generalization of the association to other stimuli [(C), top] or extinction learning [(C), bottom], where repeated exposure to blue cars diminishes the threat response, may occur. A more permanent way of diminishing the learned threat response is by modifying the original association through reconsolidation updating (D). A reminder cue may trigger the memory and destabilize it, requiring restabilization (reconsolidation) to return it to a stable state. In the course of destabilization, updates may occur in several ways, such as extinction (top), counterconditioning (middle; car associated with a positive outcome such as a wedding), or sensorimotor interference (depicted here as a Tetris game). The new information these processes provide is incorporated into the memory (extinction, counterconditioning) or depletes neural resources of reconsolidation (sensorimotor interference). Finally, threat learning interacts with processes of decision making and attitudes toward loss, risk, and ambiguity (E). For example, when facing a choice between riding in a car or on a bicycle, threat-related processes may bias the choice toward the less threatening option. The depiction of the stages of threat experience (A–E) does not mean to indicate any sequential order or independence. The stages are intertwined throughout the threat experience.

@March 29, 2023

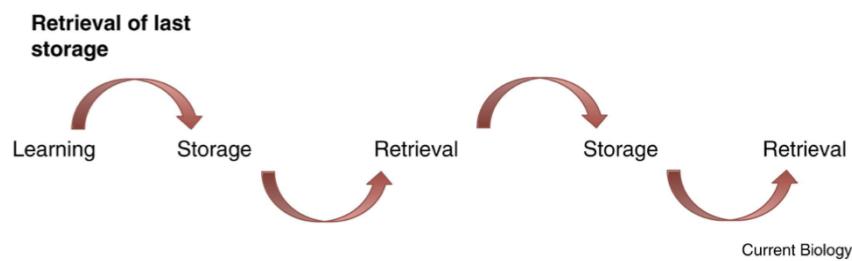


activates the CS representation, which in turn activates the US representation and triggers the fear response.

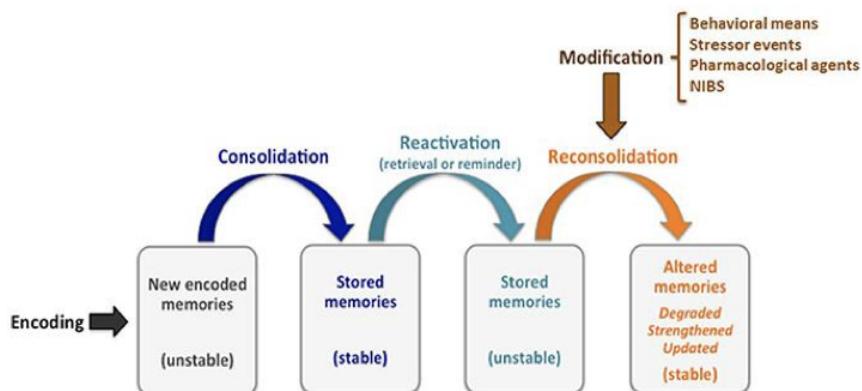
The response is illustrated as a potentiated startle (i.e. a greater amplitude startle response when the reaction is elicited in the presence of the CS relative to when it is elicited in the absence of the CS). Extinction involves the formation of a second, parallel, inhibitory association whose effect is opposite to that of the excitatory association. When this occurs, the CS representation no longer activates the US representation and no fear is triggered.

Memory re-consolidation: rather than having two parallel and opposing memory traces stored, can we have a direct modification of the acquisition memory (e.g. you learn that a car is not dangerous, after a car-induced traumatic experience)? (Kandel research on Aplysia)

Every time you retrieve an experience and recall them, the memory is susceptible to modifications and then it is stored again. You don't store exactly the same memory again ⇒ instead of an extinction trail, can we intervene directly on the acquisition trail?



This shows the idea of how memory works; after consolidation, you can reactivate the memory and you don't store exactly the same memory, but you reconsolidate and store it again (it's a cycle of retrieval-storage)



Disrupting the re-consolidation of fear memory through the use of the β -adrenergic receptor antagonist propranolol erases the behavioral expression of the fear memory 24hrs later.

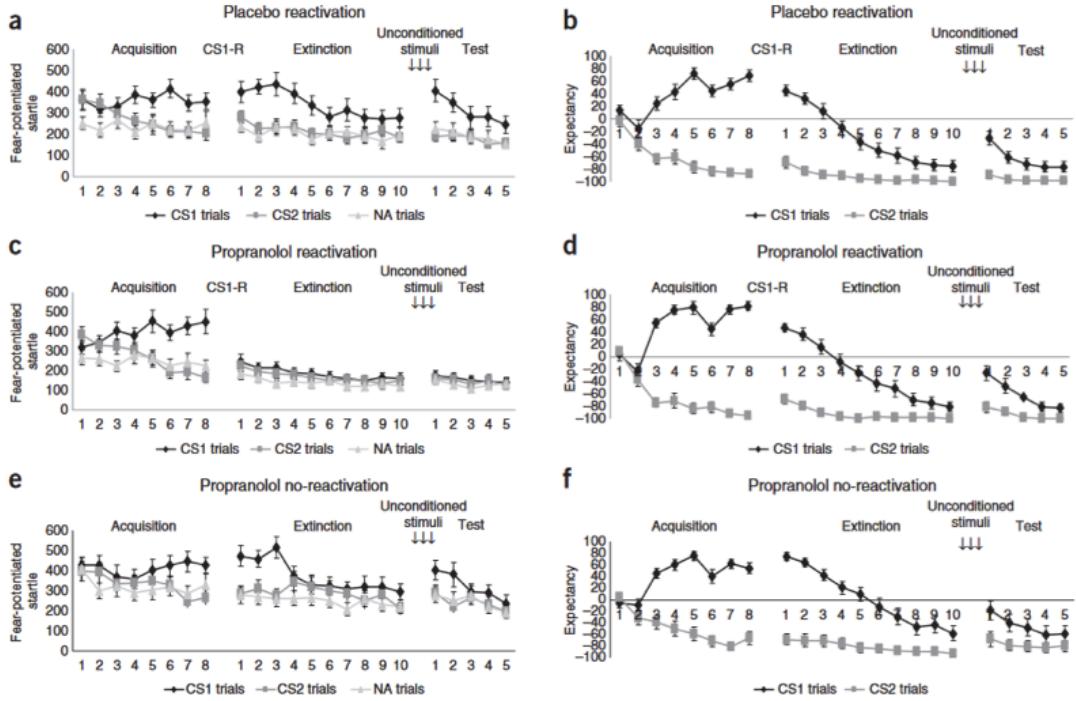


Figure 1 Propranolol disrupts the reconsolidation of a fear memory, but not declarative memory. **(a–f)** Mean startle potentiation to the fear-conditioned stimulus (CS1), the control stimulus (CS2) and noise alone (NA) trials (left) and mean expectancy scores of the unconditioned stimulus to CS1 and CS2 trials (right) during acquisition (trial 1–8), extinction (trial 1–10) and test (trial 1–5) for the placebo ($n = 20$, **a,b**), propranolol reactivation ($n = 20$, **c,d**) and propranolol without reactivation ($n = 20$, **e,f**) group. CS1⁺ refers to the fear conditioned stimulus during acquisition, CS1⁻ refers to the fear conditioned stimulus during extinction and test, CS1-R refers to the reactivation of the fear conditioned stimulus and CS2⁺ refers to the control stimulus during all phases of the experiment. Error bars represent s.e.m.

Neural Bases of Aversive Conditioning

Two different techniques used in research to investigate neural bases of different processes:

- fMRI → provides correlational evidence (whether the region is or isn't involved in the process). It measures the ratio of oxygenated to deoxygenated hemoglobin in a certain area of the brain (indirectly measure the firing of neurons = bold signal). It is based on the fact that, when neurons are firing, they need oxygen to work.

<https://www.youtube.com/watch?v=4UOeBM5BwdY>

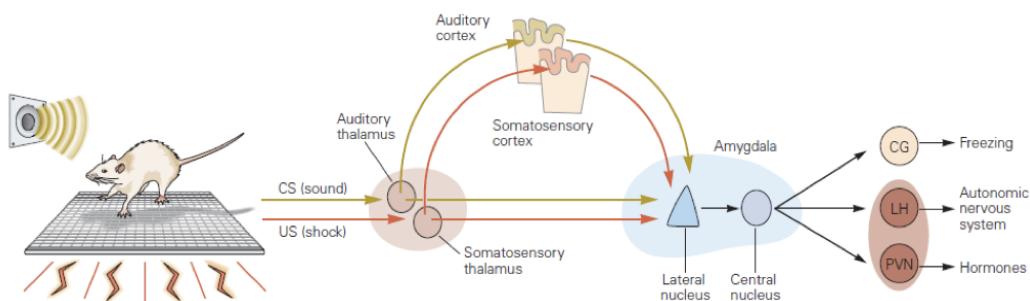
- Lesional Method → provides causal evidence (the region has a causal role in determining the response). You study the consequence of a lesion on a specific area of the brain (e.g. studies on aphasias). The lesions can be naturally occurring (trauma, tumor, stroke), surgically induced or experimentally caused (only in animals of course).

The Amygdala

The experiment wanted to test whether the amygdala participates in fear conditioning. They presented an unconditioned stimulus to a number of healthy participants. They also show a

conditioned stimulus and a “Control” Stimulus (not associated to anything) and observe the activation of the amygdala in the acquisition and extinction phase and skin conductance level difference. This though just tells that my amygdala is activated, not that it is necessary.

Another study, still from LaBar, this time with a group of participants who had a lesion to the amygdala (temporal lobectomy - 26 participants) and another group of healthy participants (23-people control group). Pavlovian fear conditioning task where the CS were two different tones, one of them associated to a US (white noise burst). Still uses skin conductance levels. The patients didn't learn to respond more to the CS+ than the Control Stimulus. The difference between CS+ and CS- was negative. They didn't acquire the conditional response to the CS+.



Redundancy and Parallel Processing \Rightarrow same information through two different pathways

The conditioned stimulus CS and unconditioned stimulus US are relayed to the lateral nucleus of the amygdala from the auditory and somatosensory regions of the thalamus and cerebral cortex. Convergence of the CS and US pathways in the lateral nucleus is believed to underlie the synaptic changes that mediate learning. The lateral nucleus communicates with the central nucleus both directly and through intra-amygdala pathways involving the basal and intercalated nuclei. The central nucleus then connects with the regions that control various motor responses.

One of the two paths is direct and the other goes through the sensory cortices. Information that relays directly to the amygdala is not processed consciously and automatically generates a response (faster). The other path makes us aware of the input (slower, detailed, enables us to consciously process the information and gives rise to voluntary response)

The Hippocampus

The hippocampus is related to declarative and explicit aspects of the memory. Same experiment as the amygdala patients (this time the lesions were hippocampal), with opposite pattern of performance: they're able to learn the autonomic conditional response, but they're not able to express “this is the stimulus that I responded to”. So, as a result, double dissociation between patients with amygdala lesions and hippocampal lesions. The

hippocampus is responsible for Declarative Memory. Opposite effect w.r.t. amygdala. They don't consciously recognize the stimulus that causes the reaction.

vmPFC

Ventral Medial PreFrontal Cortex was thought to be involved in both acquisition and extinction; this was though based on correlational evidence, without any causal evidence. The only lesional study was one of 1999, Bechara. With a small sample size, the author measured the skin conductance response over all the phases of conditioning. The data is coherent w.r.t. what we have seen before with amygdala-lesioned patients. But this doesn't show causal evidence that vmPFC has a role in acquisition and/or extinction, as the patients behaved similarly to the control group. So it is not necessary.

More recent studies performed by the lecturers. Replicate Bechara's studies. Patients with bilateral lesions to the vmPFC, 10 patients with lesions in other parts of the brain (not vmPFC, amygdala or hippocampus - damaged control) and a healthy control. The experiment was a pavlovian threat conditioning task, in which the skin conductance response and the explicit responses were measured (verbal reports). They didn't replicate the original paper. They don't form the connection between CS and US, but respond directly to the shock.

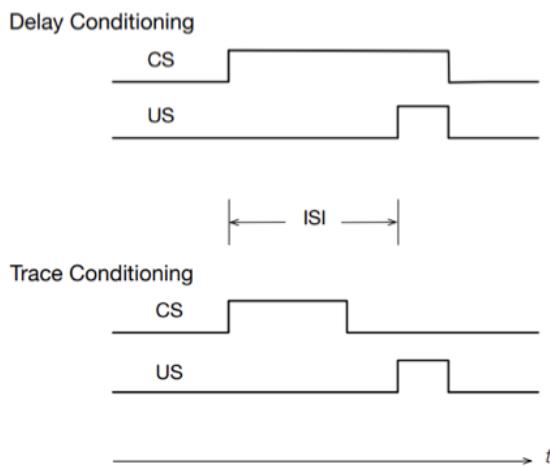
In both cases, the small sample size is a downside. So, the correlational evidence is proven, but the causal evidence is mixed. Also the measure of skin conductance could be a downside of the experiment.

What conditions enable learning? Contiguity, contingency and surprise

Is it enough to present two stimuli one after the other to actually learn? Conditioned responses are anticipatory, predictive responses (e.g. the conditioned eyelid response, the startle, the skin conductance...).

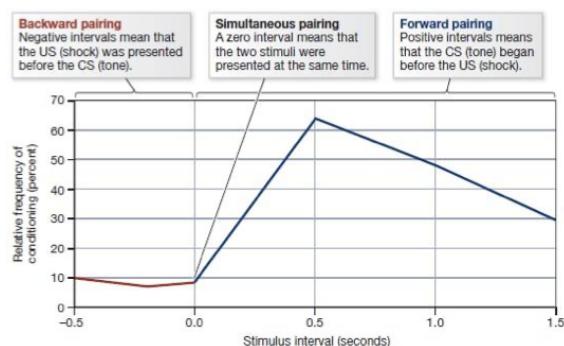
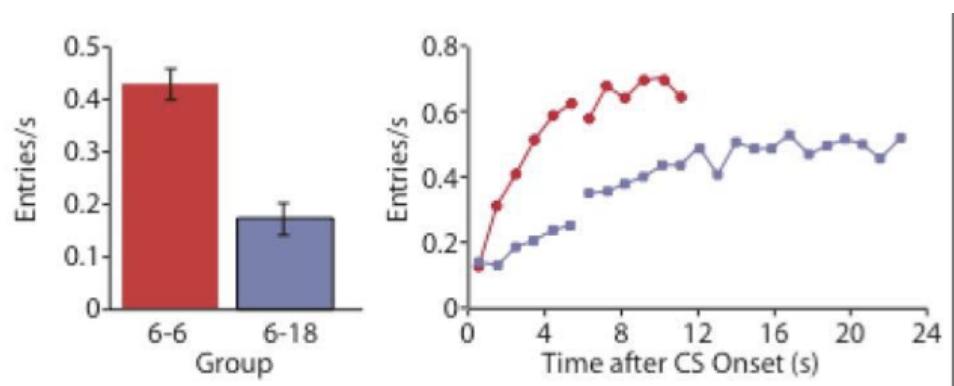
Contiguity → closeness in time between the stimulus/behavior and the outcome. Stimuli close to one another in time become associated. Initially believed to be both sufficient and necessary.

- Delay Conditioning → the CS extends throughout the inter-stimulus interval (ISI), which is the time interval between the CS onset and the US onset (often with the CS and US ending at the same time as shown)



- Trace Conditioning → there is a time interval (trace interval) between the CS offset and US onset. Assumes that a trace of the CS remains when the US arrives; therefore learning occurs through the simultaneous presence of the trace and the US. Learning does not occur across long trace intervals.

Experiment on two groups of rat, exposed to a 6s tone CS followed by a pellet delivery after different trace intervals: 6s for the first group and 18s for the second group, and recorded anticipatory head entries into the feeding hopper.



Result of studies of effectiveness of various CS-US intervals in humans. The CR was a finger withdrawal response, the CS a tone and the US an electric shock. The time between CS and US is plotted as it is, negative values indicating the CS arrived after the US. The vertical axis indicates the strength of conditioning.

Contiguous relations are important because only closely spaced presentations of the CS and US allow for their mental representation to be simultaneously active and learned about.

@April 5, 2023

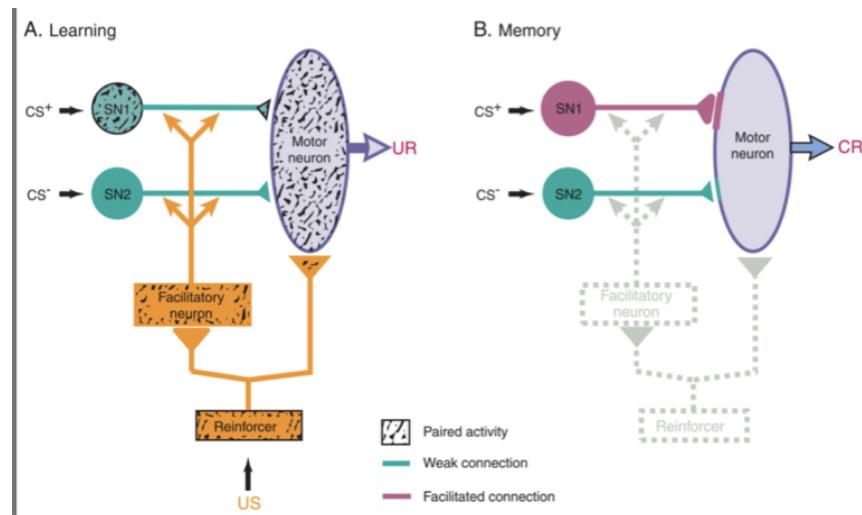
Plasticity → modification of neural connections following experience and learning, that shape the way neurons are connected. Learning at neuronal/biological level → result of changes in the strength of synaptic interaction among neurons.

Functional alterations, physiological changes are typically short-term and result in changes in the efficacy of existing synaptic connections. Anatomical alterations are typically long-term and consist of the growth of new synaptic connections between neurons.

Hebbian plasticity → contiguity learning at the neuronal level. It comes from the idea that “neurons that fire together wire together” and describes how when a cell persistently activates another cell nearby, the connection between the two cells becomes stronger.

Hebbian Plasticity

Proposed mechanism to explain synaptic plasticity (Neurons that fire together wire together).



How the circuit is structured and how it functions in the Aplysia

(A) Learning → Activity in one sensory neuron SN1 is paired (CS1) with the reinforcing stimulus US. Activity in SN2 is unpaired with the US. The US itself acts 2-fold by activating the motor neuron directly, thus producing the unconditioned response (UR), and by activating a modulatory system (facilitatory neuron) that nonspecifically enhances the synaptic strength of both sensory neurons. The paired activity in SN1 results in a selective amplification of the facilitation caused by the US.

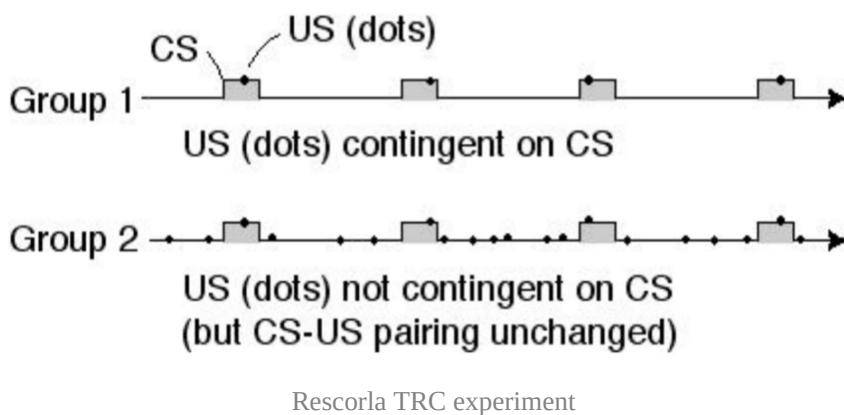
(B) Memory → as a result of the paired activity, the synaptic strength in the SN1 is enhanced, which increases the probability to activate the motor neuron and to produce the conditioned response (CR). Because activity in SN2 was unpaired with the US, the connection of SN2 is not specifically enhanced.

Touching the siphon automatically triggers the response (close the gill), so it is our US. The information is relayed directly to the motor neuron, which also has an indirect effect from the sensory neurons activated by the interneurons and interventions (CS+ is the somatosensory sensation of touching the tail - touching the tail while using the US will provoke the reaction - CS- is not providing stimulation - acts as control). The sensory and motor neuron are activated together → strengthened synaptic connection. Even when there's no US anymore, due to the synaptic plasticity, the pairing stays there and works. The CS- is the control signal.

https://www.youtube.com/watch?v=-wakk5Iwqj8&ab_channel=DanielKochli

Truly Random Control Paradigm: why contiguity is not sufficient

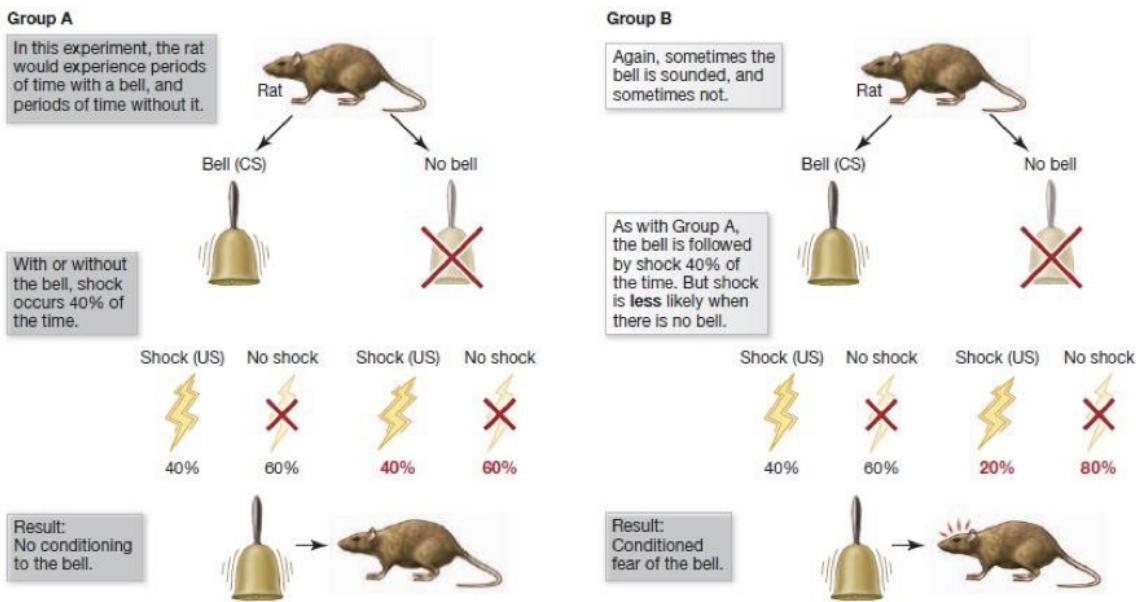
Contiguity is not sufficient, which is demonstrated by the Truly Random Control Paradigm, by Rescorla (1968) which shows that temporal contiguity between a conditioned stimulus and an unconditioned stimulus is not sufficient to lead to learning.



Two groups of beings. Gray marks = CS, dots = US. In group 1 the CS and US happen every time together. In group 2, the CS and US are still close together (it's actually the same as group 1 in this sense), but the US happens many more times outside the CS range (the CS is not a good indicator of when the US will arrive)

This is the **Truly Random Control Paradigm** (Rescorla) → temporal contiguity between CS and US is not sufficient to lead to learning

Contingency = the probability that given the CS the US should appear should be greater than the probability of it happening if the CS isn't there.



7.11 The effect of contingency on classical conditioning For both groups, there's only a 40% chance that bells will be followed by shock. However, for Group B, shock is less likely when no bell is sounded, and, for this group, the bell becomes a fearful stimulus.

In group A, the bell is not predictive enough (the rat has a 40% shock chance anyway). But in group B, the shock is more likely with the bell than without the bell. So, it's also a matter of learning the causal relationship between the two events

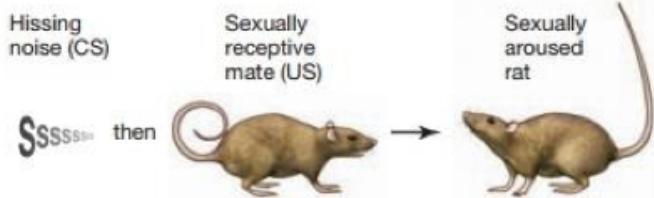
Contingency → causal relationship between stimulus/behavior and outcome. When one stimulus is dependent on the other, they will become associated (predictive value is critical).

The Surprise factor

But you also need surprise.

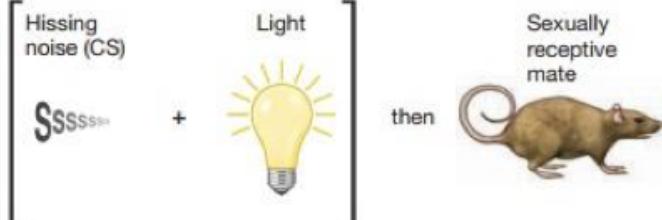
Stage 1

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



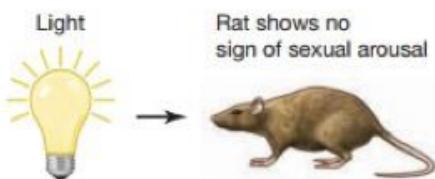
Stage 2

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.



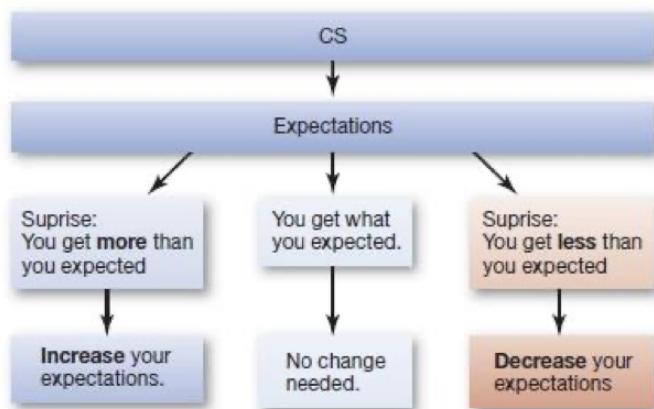
Stage 3

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



Blocking phenomenon: no learning occurs to a stimulus if it is combined with a previously conditioned stimulus during conditioning trials.

Learning is regulated by prediction error (quantitative discrepancy between the expected outcome when the cue is presented and the experienced outcome). Functions as teaching signals to update expectations and reduce following prediction errors.



Learning theories have posited that conditioning is not instructed by a simple sensory representation of the outcome, but instead by an error signal measuring the difference between the outcome actually present and that expected, which is the prediction error.

The prediction error is necessary for associative learning: it dictates variations in the effectiveness of the US in supporting learning. If the difference is large, predictions did not match observations and there is a need for more learning to update those predictions; else, if predictions match observations, there is no prediction error and no learning occurs.

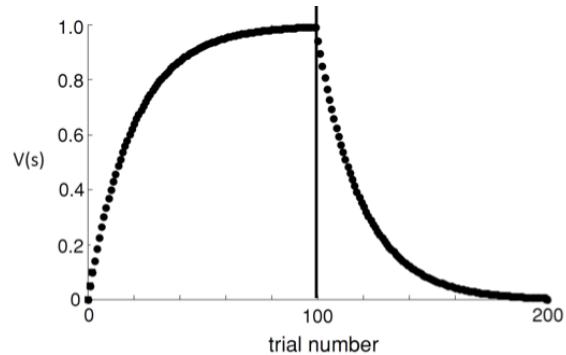
Computational Models

Mathematical Formalization of learning theories

Rescorla-Wagner model

Formalize the idea that learning happens with prediction error (Error-Driven Learning). R is the value of US. 1 if delivered, 0 if omitted. V is the expectancy. We observe that overall $V_{t+1} = V_t + \alpha\delta_t$, where δ_t is the prediction error at trial t , given by $R_t - V_t$. α is a learning rate parameter, which determines the size of the update step.

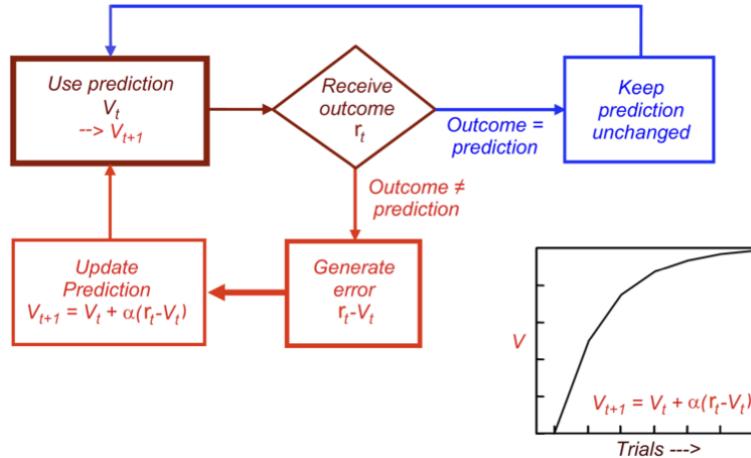
Learning is proportional to the prediction error, which is larger at the start of training when reward is unexpected, and gets smaller (eventually 0), as training progresses and rewards are fully predicted.



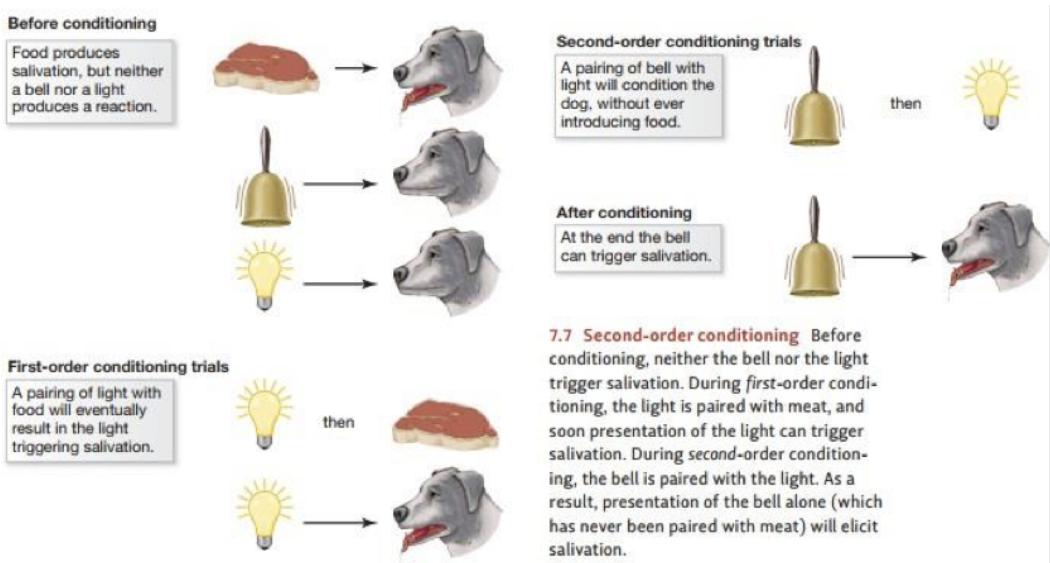
Prototypical acquisition and extinction learning curves for Pavlovian conditioning. The filled circles show the time evolution of the value V of a CS over 200 trials, where in the first 100, a reward was paired with the CS (acquisition), later not (extinction)

- δ is positive (more than expected) and larger at the start of acquisition, when reward delivery is unexpected. It progressively gets smaller, as acquisition progresses and reward delivery is fully predicted
- δ is negative (less than expected) and larger at the start of extinction, when reward omission is unexpected. It eventually tends to 0 when extinction progresses and reward omission is fully predicted.

The curve is similar to what we observed before in acquisition and extinction.



Experimental result, but didn't explain all of them. Trial-level model → deals with how associative strengths change from trial to trial, but without considering any details on what happens within and between specific trials.



RW model predicts no learning because there is no US in the second-order conditioning trials, but the animal indeed learns.

Temporal Difference Model (Sutton-Barton)

The trial is not an entity, but a fragment of the experience of interacting. Real time model. t represents a time-step within the trial. It provides a prediction, for each time t in the trial during which a CS is presented, of the total future reward that will be gained in the trial from time t to the end of the trial. The prediction error compares the predicted value at time t to the predicted value at $t - 1$.

Prediction errors in CM are signed values used to update the predicted value → the value signal produced by the reward transfers back to the events that reliably precede reward

delivery → rewarding value

Signed difference between the expected and delivered outcome

- Used to update predicted value
- The value signal produced by the reward itself transfers back to events that reliably precede reward delivery (e.g. CSs)
- Thus, the rewarding value transfers from the reward to the CS that predicts reward

In this manner, reinforcement learning algorithms explicitly state that the quantitative value inherent in reward transfers back to the antecedent cue predicting its delivery. That is, the predictive cue becomes endowed with the scalar value of the reward.

We make predictions all the time (not always correctly, though)

The Reward Prediction Error Hypothesis of Dopamine Neurons

Prediction error is the signed difference between the expected and delivered outcome; it functions as a teaching signal to update expectations and reduce following prediction errors. It is used to update the predicted value: the value signal produced by the reward itself transfers back to events that reliably precede reward delivery, thus, the rewarding value transfers from the reward to the CS that predicts reward.

How could we implement prediction error in neurons? Use synaptic plasticity to impose changes in the amount of neurotransmitter that is released or neuromodulator that gets released (effect is other than direct neuron excitation or inhibition). This is a way through which firing rate can change.

The Reward Prediction Error Hypothesis of Dopamine Neuron Activity

Dopamine is a neuromodulator that plays an essential role in many processes (e.g. motivation, learning, decision making, addiction...)

https://youtu.be/Wa8_nLwQIpg

Three major pathways

1. Nigrostriatal pathway → critical role in the production of movement as part of the basal ganglia motor loop. Originates in the substantia nigra pars compacta (SNc) and projects primarily to the caudate-putamen.
2. Mesolimbic pathway → originates in the VTA, projects to the nucleus accumbens, septum, amygdala, hippocampus.
3. Mesocortical pathway → originates in the VTA, projects to the medial prefrontal, cingulate, orbitofrontal and perirhinal cortex.

These 2 pathways are important for motivational functions (reinforcement learning).

The Reward Prediction Error Hypothesis resulted from the convergence of computational reinforcement learning and results of neuroscience experiments. The modulation of synaptic plasticity via the neuromodulator dopamine is a plausible mechanism for how the brain might implement learning algorithms. There is strong evidence that the dopaminergic system is the major neural substrate of reward and reinforcement for both natural rewards and addictive drugs.

https://youtu.be/8n1-lv1_wCY

Schulz, Dolan, Dayan, The Brain Prize, 2017

The early view, from 1954, was that dopamine neurons broadcast a reward signal: activation by unconditioned rewarding and aversive stimuli. About 75% of dopamine neurons show phasic activation when animals touch a small piece of hidden food, or when drops of liquid are delivered to the mouth outside of any task. They do not discriminate between different food or liquid rewards; however, their responses distinguish rewards from non-reward objects. Only 14% of dopamine neurons show phasic activations when aversive stimuli are administered (e.g. pain pinch, electrical shock). However, they are not entirely insensitive to aversive stimuli, as they show depressions or activations with slower time courses following pain pinch. The phasic activations of dopamine neurons report preferentially environmental events with rewarding value, whereas aversive events may be signaled primarily with a slower time course.

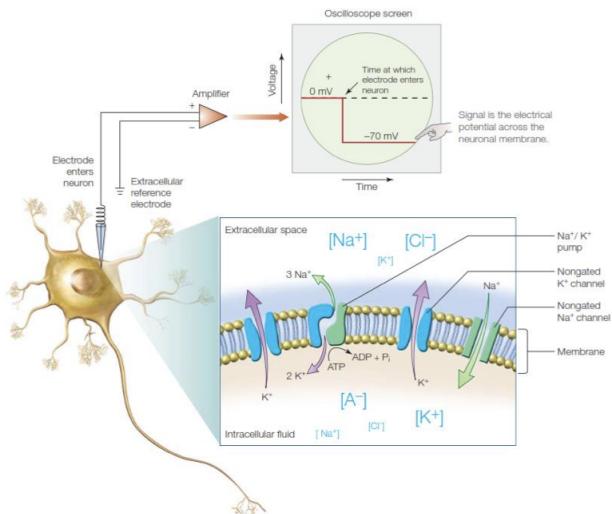
Dopamine neurons show phasic changes in firing rate when predictions do not meet reality:

1. Increase firing rate when a reward is unexpectedly delivered or is better than expected
2. Suppress firing rate when a reward is unexpectedly omitted or is worse than expected

Phasic responses of dopamine neurons signal reward prediction errors, not the reward itself.

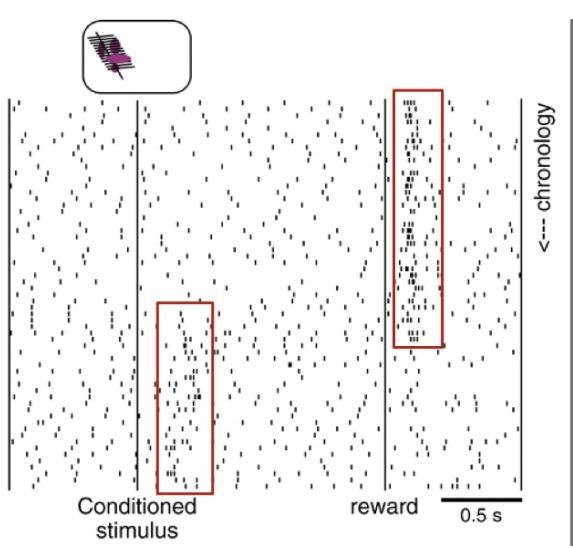
Measuring neuronal signalling

Idealized neuron shown with intracellular recording electrode penetrating the neuron. The electrode measures the difference between the voltage inside versus outside the neuron and this difference is amplified and displayed on an oscilloscope screen (top). The oscilloscope screen shows voltage over time, and shows that prior to the electrode entering the neuron, voltage between the electrode and the extracellular reference electrode is zero, but when the electrode is pushed into the neuron, the difference becomes -70 mV , which is the resting membrane potential.



Before learning, an unexpected reward occurs, which means a positive prediction error. Dopamine neuron firing is increased following reward. After learning

- the CS predicts reward, and the reward occurs (so, there's no prediction error): the dopamine neuron firing is increased following the CS but not following the reward. So dopamine \neq reward signal.
- the CS predicts a reward, but it does not occur, which means that I have a negative prediction error: the dopamine neuron firing is increased following the CS, but decreased following the omitted reward, exactly at the time when reward was expected.

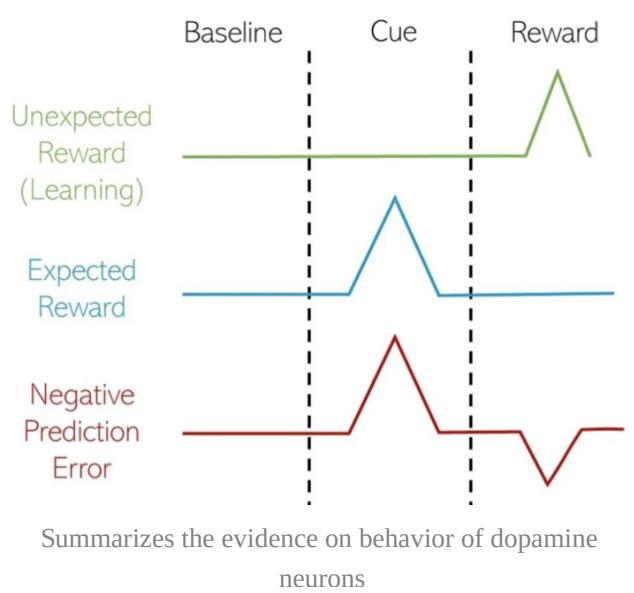


Each line of dots represents a trial, each dot represents the time of the discharge of the dopamine neuron, the vertical lines indicate the time of the stimulus and juice reward, and the picture above the raster shows the visual conditioned stimulus presented to the monkey on a computer screen.

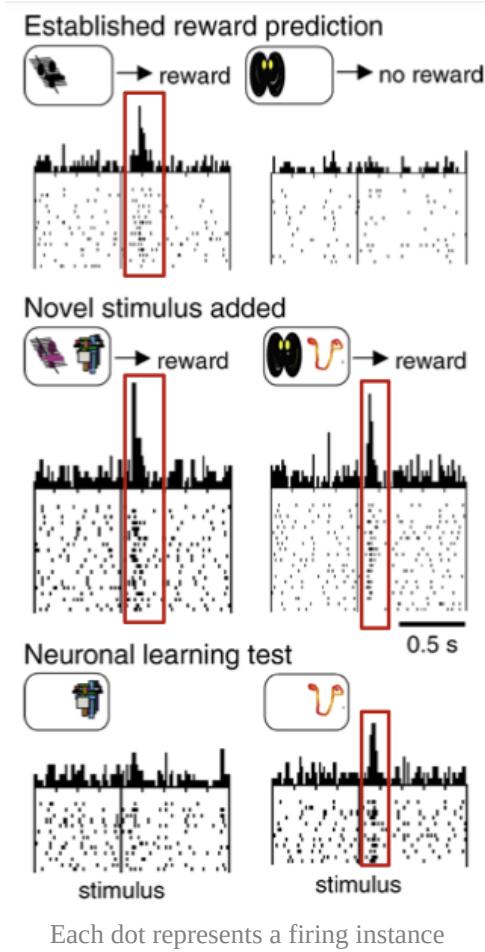
Chronology of trials is from top to bottom. The top trial shows the activity of the neuron while the animal saw the stimulus for the first time in its life, whereas it had previous experience with the liquid reward.

@April 12, 2023

Unexpected rewards increase the activity of dopamine neurons, acting

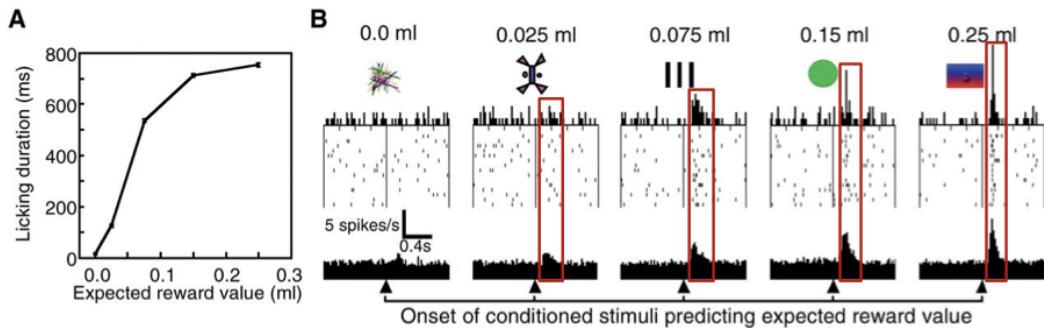


as positive feedback signals for the brain regions associated with the preceding stimulus/behavior. As learning takes place, the timing of the activity will shift until it occurs upon the cue alone, with the expected reward having no additional effect. Should the expected reward not be received, dopamine activity drops, sending a negative feedback signal that weakens the positive association.

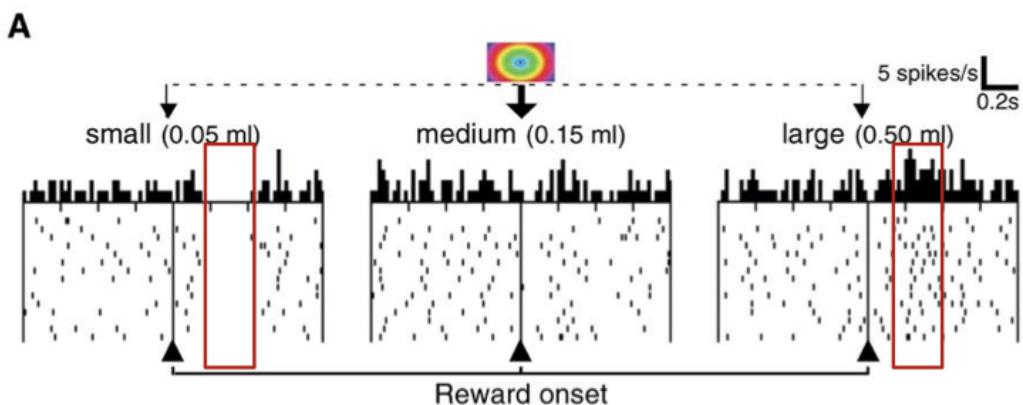


Blocking tells us that we only learn about something we didn't know/predict before, aka we need surprise. Now, we want to test if dopamine neurons conform to blocking, using an experiment based on compound stimuli; the experimenter just shows the second stimulus of the compound. If blocking happens, we should expect no learning of the second stimulus in the first case, while the second case should show a form of learning. The results show that dopamine neurons conform to blocking.

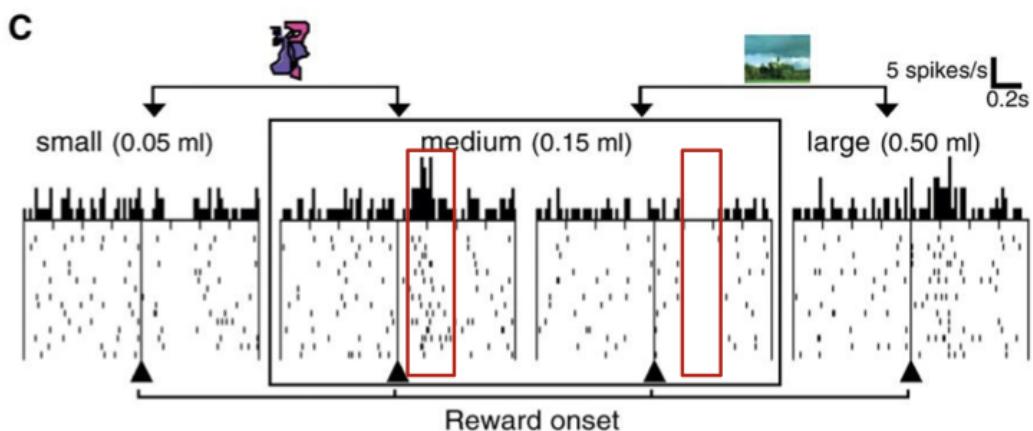
We want to inspect also if the signal is modulated based on the quantity



When the condition stimulus doesn't predict anything there isn't much of a change, and it increases the more stimulus is provided.



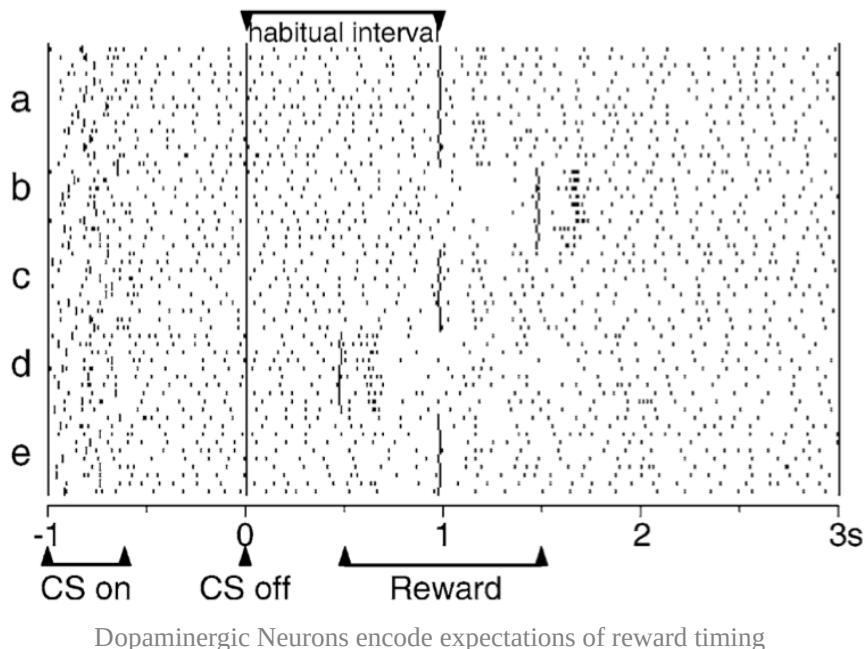
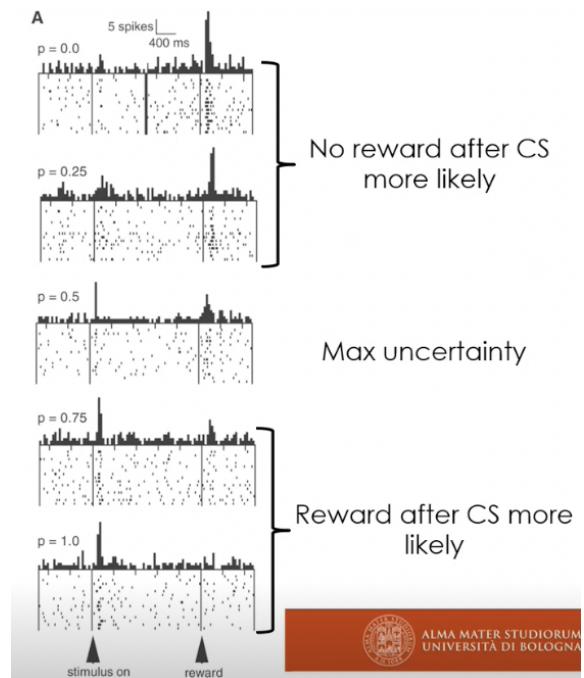
When we modify in a remarkable way the quantity of stimulus, you find a modification in the firing



Responses of a single neuron to three liquid volumes, delivered in the context of two different predictions. One stimulus predicted small or medium volume with equal probability, whereas another stimulus predicted medium or large volume. The medium volume activated the neuron in one context, but suppressed activity in the other:

- Dopamine neurons process reward magnitude relative to a predicted magnitude
- A reward outcome that is positive on an absolute scale can nonetheless suppress the activity of dopamine neurons.

Phasic activation of dopamine neurons vary monotonically with reward probability. The experiment shows the response of the dopamine neuron when the stimulus is presented. The reward is given with different probability.



Temporal prediction error: firing is depressed exactly at the time of the usual occurrence of reward when a predicted reward is omitted. The depression occurs even in the absence of any stimuli at the time of the omitted reward. The depression does not constitute a neuronal

response to a stimulus, but reflects an expectation process based on an internal clock tracking the precise time of predicted reward.

Dopamine neurons encode also the omission of reward. The prediction error is also a temporary prediction

To sum up the experiments seen so far:

Dopamine neurons do not broadcast a reward signal, but indeed broadcast a prediction error signal. Dopaminergic neurons exhibit changes in phasic response to:

- an unexpected reward → discriminate between reward & no reward, reward magnitude (more or less) in a relative way, reward probability (more or less likely), reward timing, in a signed manner: increase firing to unexpected delivery and pause firing to unexpected omission
- transfer back to a cue which predicts reward occurrence, enabling associative learning

So Dopamine is not simply the feel-good chemical, but it makes us feel good when something valuable unexpectedly happened, make events that unexpectedly predicted rewards relevant, while disregarding those events that didn't and teaches us where to find things we need or like by conveying PEs.

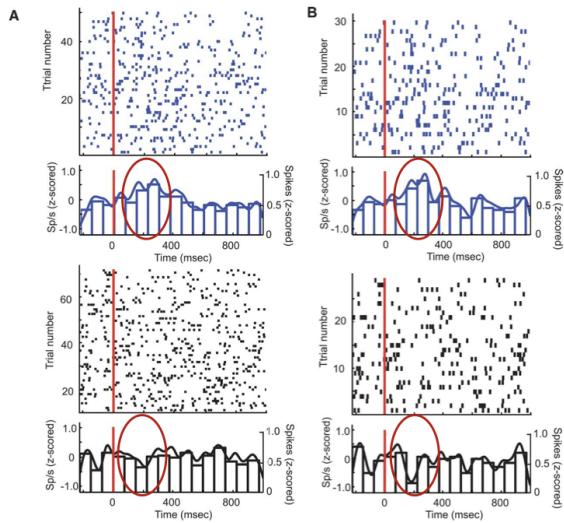
Human Studies

Human studies also demonstrate reward prediction error signals in the human brain.

Electrophysiological correlational evidence

Microelectrode recordings during deep brain stimulation surgery to study neuronal activity in the human substantia nigra while patients with Parkinson's disease engaged in a probabilistic instrumental learning task motivated by virtual financial rewards. The participants are presented with two decks of cards in a computer screen. They are instructed to repeatedly draw cards from either deck to determine which deck yields the higher reward probability.

(A) Spike raster for a single experiment from one participant. Individual spike activity recorded from SN for trials during positive (blue) and negative (black) feedback is shown for each trial as a function of time. Below each spike raster is the average z-scored continuous-time firing



rate and histogram. The red vertical line indicates feedback onset.

(B) Individual spike activity from the same cell shown in A for trials in response to unexpected gains (blue) and losses (black) as a function of time.

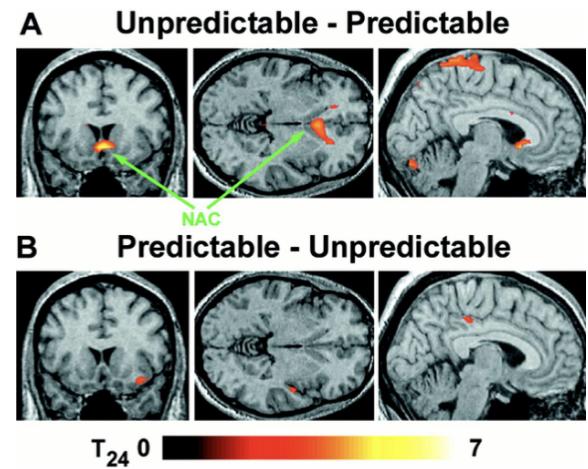
Raw spike count increased in response to positive feedback and decreased in response to negative feedback during this interval. The difference in activity between responses to unexpected gains and losses was clearer than the difference between positive and negative feedback.

Neuroimaging correlational evidence

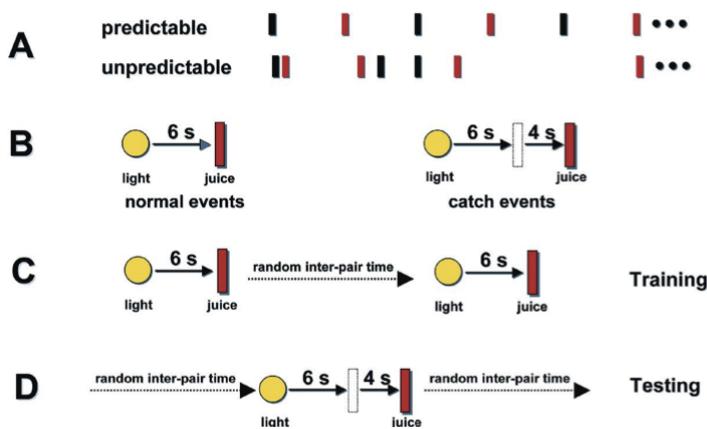
fMRI experiments to obtain neuroimaging correlational evidence. A 2x2 factorial design was used, with factors of preference (juice or water) and predictability. Subjects received 0.8 ml boluses of juice and water in either a predictable or unpredictable sequence.

Reward-related regions had a greater BOLD response to the unpredictable stimuli.

Planes centered at (0,4,-4) show that bilateral nucleus accumbens/ventral striatum (NAC) and bilateral superior parietal cortex were more active in the unpredictable condition, while a small region in the right superior temporal gyrus was relatively more activated by the predictable stimuli



In the previous experiment, juice and water were delivered to subjects in two separate sequences. Stimuli derived



during the unpredictable sequence were associated with greater changes in brain activity in the ventral striatum compared with stimuli delivered during the predictable sequence.

What follows is a follow-up of this experiment.

However, stimuli can be unpredictable in character (what arrives next), in time (when) and in quantity (how much). The experiment sought to separate the effects of temporal prediction errors only. Subjects were trained to expect juice at a fixed time following a flash of light (normal events) and then changes in brain response were probed when juice was delivered at an unexpected time (catch events)

Normal events consisted of brief flashes of a yellow light centered in their visual field and orally delivered fruit juice. The time between individual events was randomly selected.

After 49 consecutive light-juice pairings, several catch events were randomly inserted among normal events. For the catch events the time of juice delivery was extended.

Comparing the brain response to juice delivered during catch and trained events reveals the effect of predictability on the induced brain response. Unpredictable juice delivery is associated with significantly greater activity in the left putamen. Combining the average response amplitude across subjects shows that predicted juice delivery induces essentially no change in fMRI signal, whereas unpredicted juice delivery induces a significantly positive change.

Comparing the brain response at 6s following the light in catch events vs a period of no juice delivery in normal events (10s following light) reveals brain regions that correlate with the negative prediction error. The failure of juice delivery during catch events correlated with decreased activity selectively in the left putamen. Significant decrement in BOLD signal following the absence of juice delivery at expected times.

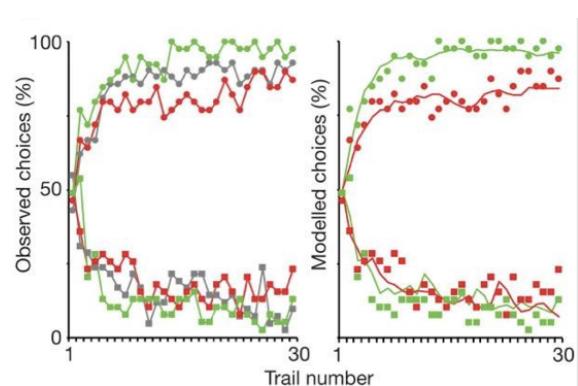
So far, we still haven't showed direct causal evidence of a link with dopamine and of an effect on behavior.

Behavioral & Pharmacological Causal Evidence

During instrumental learning, the behavioral and neural effects of drugs enhancing dopaminergic functions, reducing dopaminergic functions and placebo (for control) was

assessed in groups of healthy subjects. The task consisted in selecting either the upper or lower of two abstract visual stimuli presented and subsequently observed the outcome.

Each pair of stimuli was associated with a certain probability of gains (assess the effects of the drugs on the ability to learn from rewards), loss (assess the effects of the drugs on the ability to learn from punishments) and neutral (control).



Observed behavioral choices for placebo (gray), L-DOPA (green) and haloperidol (red) and modeled behavioral choices for L-DOPA and haloperidol

The learning curves depict the proportion of subjects that chose the correct stimulus in the gain condition and the incorrect stimulus in the loss condition.

In the right graph, the learning curves represent the probabilities predicted by the computational model.

A standard algorithm of action-value learning was then fitted to the observed behavior. Outcome prediction errors estimated were then used as a statistical regressor in the imaging data. Brain activity correlated with prediction errors derived from the computational model. Reward prediction errors correlated with activity in the left posterior putamen, left ventral striatum. Punishment prediction errors correlated with activity in the right anterior insula.

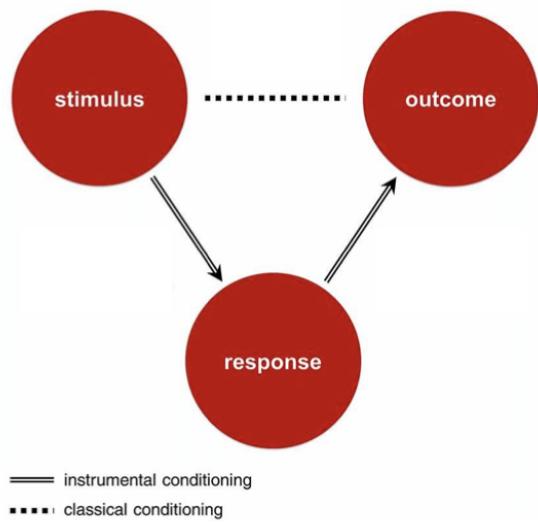
So, in human studies, the signals reflect the dopamine prediction error response, which occurs in striatal and frontal dopamine terminal areas rather than in midbrain cell body regions, presumably because it reflects summed post-synaptic potentials.

Linking motivation/value to action: the Actor-Critic model

Dopamine reward prediction errors are used to learn the values of states (Pavlovian), state-action (Instrumental) or both. These values are then used to select optimal actions.

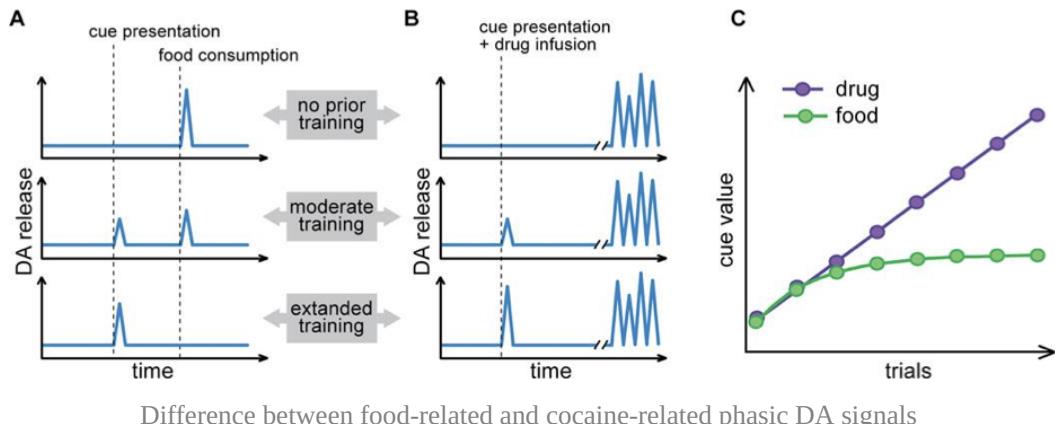
The cortex represents the current state and the basal ganglia implement two computational models:

1. the critic → learns state values (stimulus/outcome associations) and may be implemented in the ventral



striatum and possibly in the amygdala and orbitofrontal cortex (OFC)

- the actor → learns stimulus-response associations and may be implemented in the dorsal striatum



Instrumental Learning

Instrumental Conditioning involves associating (it is a form of Associative Learning) an action/behavior with an outcome. Instrumental and Pavlovian Conditioning are strictly related to each other as if we can predict what situations are associated with rewards we can try to bring those about through our actions - aka, by using Pavlovian conditioning we can learn the value of a stimulus and we can use this knowledge to shape our behavior according to it. It was discovered by Edgar Thorndike (Thorndike's Law of Effect: "of several responses made to the same situation, those which are accompanied by satisfaction will be more likely to recur; instead, actions and behaviors that are followed by discomfort will be less likely to occur"). Skinner studied it through the use of a "Skinner Box": there's a lever you can press to obtain food.

Shaping (e.g. children education) → an existing response is gradually changed across successive trials towards a desired target behavior by reinforcing exact segments of behavior.

	Delivery of outcome	Omission of outcome
Appetitive outcome	Positive reinforcement	Negative punishment
Aversive outcome	Positive punishment	Negative reinforcement

Positive reinforcement → delivery of rewarding outcome increases the probability of emitting the action

Positive punishment → delivery of aversive outcome decreases the probability of emitting the action

Negative reinforcement → omission of aversive outcome increases the probability of emitting the action

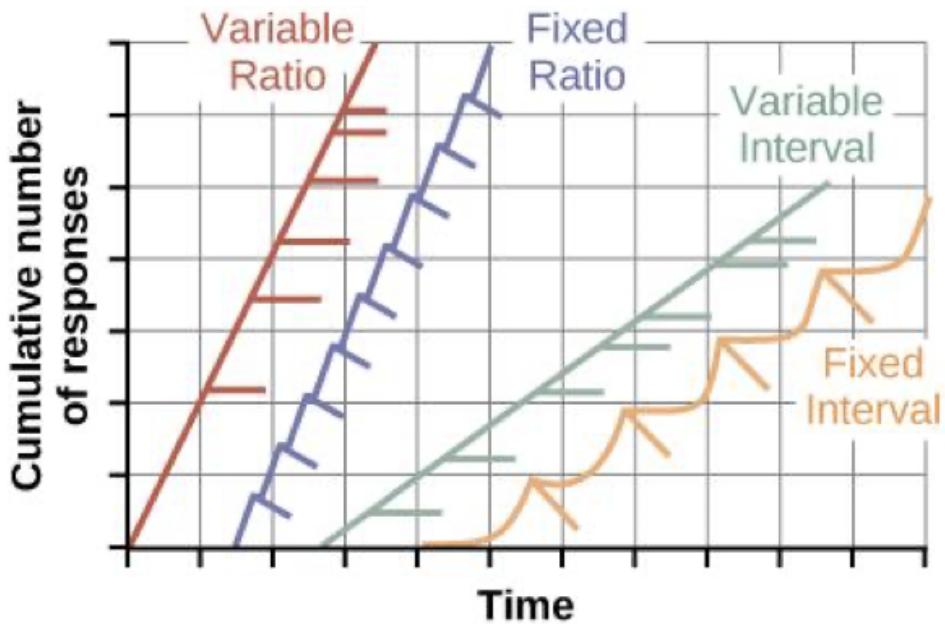
Negative punishment → omission of rewarding outcome decreases the probability of emitting the action

The nature of the outcome shapes the behavior

To make an example, Social Media has a reinforcement schedule.

Also the frequency of the outcome shapes behavior:

- Continuous schedule → the desired behavior is followed by the outcome every single time it occurs; it is most effective when trying to teach a new behavior
- Partial Schedule → the desired behavior is followed by the outcome only part of the time it occurs; in this case, behaviors are acquired more slowly, but the response is more resistant to extinction. Four different partial schedule with different effectiveness:
 - Fixed-Ratio → outcome is delivered only after a specified number of responses. Produces a high, steady rate of responding with only a brief pause after the delivery of the outcome
 - Variable-Ratio → outcome is delivered after an unpredictable number of responses. Creates a high steady rate of responding
 - Fixed-Interval → outcome is delivered after a specified interval of time. High amount of responding near the end of the interval, but slower responding immediately after the delivery of the outcome
 - Variable-Interval → outcome is delivered an unpredictable interval of time. Slow, steady rate of response



<https://www.youtube.com/watch?v=GLx5yl0sxeM>

Goal-Directed/Model-Based vs Habitual/Model-Free Behavior

@April 19, 2023

When we have a maze we have to navigate, we can learn how to do it forming a cognitive map or learn through a stimulus response strategy. Studies using rats: put a rat at one end of a T-Maze and it needs to reach the reward. Now that it has learned how to reach the reward, you make it start from another branch of the T-Maze and see if it succeeds or fails in locating the reward to determine which approach it used → if it fails it was response strategy, else it was a place strategy. There are cues to determine in which branch the animal is. Learn about places more readily using a place strategy, but with extensive training they shift to stimulus response.

Again, although we're talking about spatial navigation, we're still in a maximizing-reward scenario. We act to produce outcomes that are desirable or to avoid those that are harmful or aversive (Thorndike). How flexible are we, though? Are our choices directed by the goal or automatically triggered based on our past experience?

Generation 0: Cognitive Maps vs Stimulus-Response

How do the rat solve the maze? Two different theories:

- Stimulus-Response theories → derived from behaviorism. Solving the maze is a matter of individual stimulus-response one-to-one connections. Learning depends on the strengthening of some connections and the weakening of others
- Field Theories → solving the maze is a matter of creating a mental/cognitive map that includes multiple sets of connections. The mental map guides the responses the animal will perform and acts as a representational template.

The way in which the two theories were tested was by using the so-called Tolman's Maze to inspect the strategy. You have the entrance at one end and food at the exit. There are doors and curtains to limit the knowledge of the rat and force it into decisions (curtains to limit the view of the corridors, doors at crossroads to prevent from seeing most of the junctions). It is a formidable task.

- Experiment 1:

Two groups of rats, they both have to find their way out. One will be rewarded with food, the other one with nothing. The rewarded group solves the maze faster. There's dopamine prediction error involved. Is it true, though, that we don't learn anything without a reward (as the S-R theories argue)?

- Experiment 2:

Add a new group of hungry rats, that is rewarded only after a number of days (11) have passed. Performance grows to match (and seemingly surpass) the group of all-along-rewarded hungry rats.

So, technically, the hungry-non-rewarded rats are still learning, regardless of the reward, but they're not keen on showing it as they get nothing from it.

Latent Learning is learning which is not shown behaviorally until there's a motivation to show it and it occurs without any reinforcement provided or associations learned. Again, this study suggest that the rats are learning to have a Cognitive Map of the maze, and not exploring to trial-and-error.

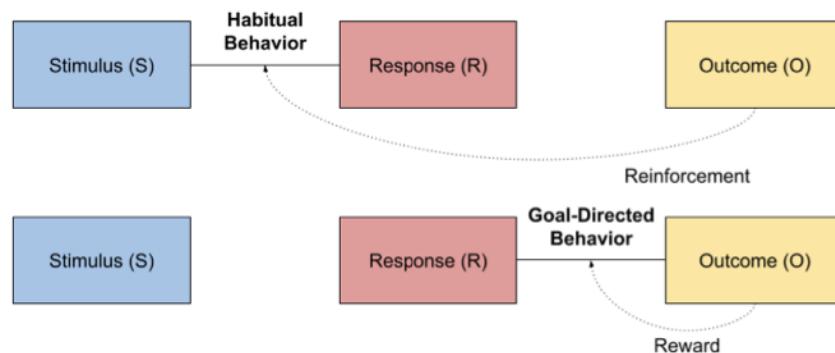
Conclusions

The experiment challenged the constraints of behaviorism, which stated that processes must be directly observable and that learning was the direct consequence of conditioning to stimuli; it also challenged the prevailing SR view of learning and behavior, which corresponds to the simplest model-free way of learning policies. Conditioning actually involves more than the simple formation of associations between sets of stimuli or between responses and reinforcers. It includes learning and representing other facets of the total

behavioral context. The phenomenon of latent learning shows you can learn also something which is not yet observable (the rats were learning from day 01, but that couldn't be seen).

Generation 0 studies established a dichotomy between decision behavior controlled by a cognitive map and by S-R associations.

Generation 1: Goal-Directed vs Habitual Actions



Operationalized the evidence from the spatial navigation context and abstracted these ideas of use cognitive maps to choose appropriate actions in non-spatial domains. What strategies we use in different tasks? Goal-Directed behavior/actions contrasted to Habitual Behavior (SR). Still animal studies + neural bases of the behavior.

How do we strengthen the relationship between Stimulus, Response and Outcome. What does the outcome reinforce?

An action is goal directed if:

1. You have knowledge of the relationship between the action and the consequences: this means we need a Response-Outcome control
2. The outcome should be motivationally relevant or desirable at the moment of choice

Goal-directed behavior involves active deliberation, with high computational cost. Adapts to the environment → if at some point you realize the action is not useful anymore (one of the two conditions falls down) you can just stop.

Habitual actions are made automatically, just because they've been rewarded in the past, they're not influenced by the current value or desire of the outcome. They don't stop, even when the outcome is undesired. In general, habitual behavior is automatic, with low computational cost and is inflexible to changing of environmental contingencies.

Testing the learning strategy (goal-directed behavior):

Testing the learning strategy (habitual behavior):

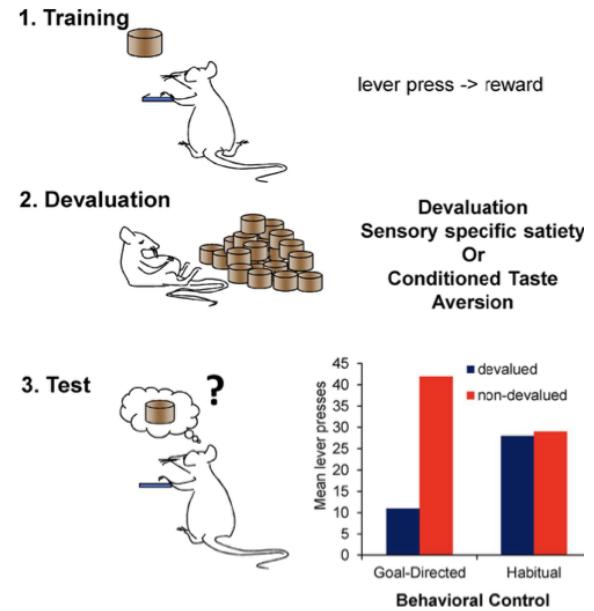
1. Training → instrumental learning
(press the lever, get a reward)
2. Post-learning manipulation:
 - a. Reinforcer Devaluation → give so much food to the rat so that the reward has no more value or associate the reward to a negative outcome
 - b. Contingency Degradation → we teach the rat that pressing the lever will not anymore bring a reward

If the behavior is goal-directed, we have so broken one of the two characteristics that make the behavior goal-directed.

3. Testing under extinction → if the action is performed less, the behavior is goal-directed, else it is habitual

1. Training → In this case, the training session based on instrumental learning is extensive, to reach a condition of overtraining

The process then is the same as before, with the same implications.



Studies useful also for the neural bases of this. Two different parts of the striatum were supporting the two behaviors. Dorsomedial striatum supports goal-directed behavior, Dorsolateral striatum supports habitual behavior.

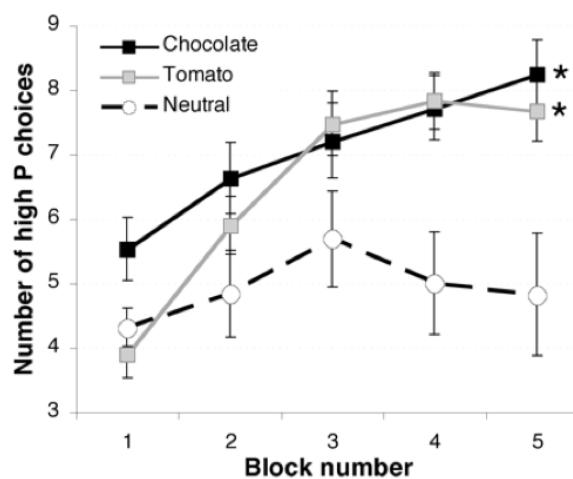
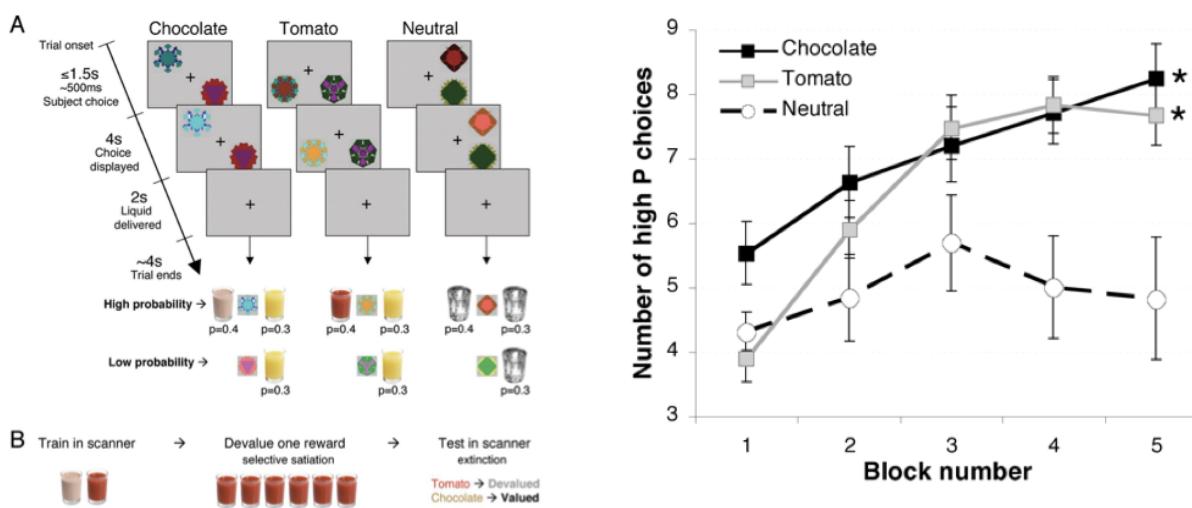
Conclusions

There's need for overtraining to make behavior habitual. So the behavior is initially goal-dependent and becomes habitual through experience. The shift to goal-dependent to habitual is an important mechanism in underlining addictive behavior. There's therefore an interdependency between goal-directed and habitual systems, which may act simultaneously and competitive. Striatum links motivation and actions.

Generation 2: Actions and Habits in the Human Brain

Focus on the human brain, looking for behavioral and their neural bases, applying paradigms that were successful with animals. Use of fMRI in order to investigate the neural bases of goal directed actions and habitual actions.

Method: different human subjects are placed into a scanner and had to learn associations between actions and rewards (Instrumental Conditioning task). Rewards were: milk+chocolate or tomato juice (they're not Italians). Each time on the screen they saw two different pictures, the one chosen was more linked to one reward than the other. Control pair had no reward. Once you come out of the scanner, you get one of the two rewards until full (devaluation) and then re-do the process performed under extinction (no reward is given). Test what happens when you don't have rewards vs you have. fMRI was recorded at train and test to examine brain areas responding during action selection, looking for areas that showed sensitivity to the change in value of the associated outcomes; such areas would be candidate regions for implementing goal-directed behavior in humans.

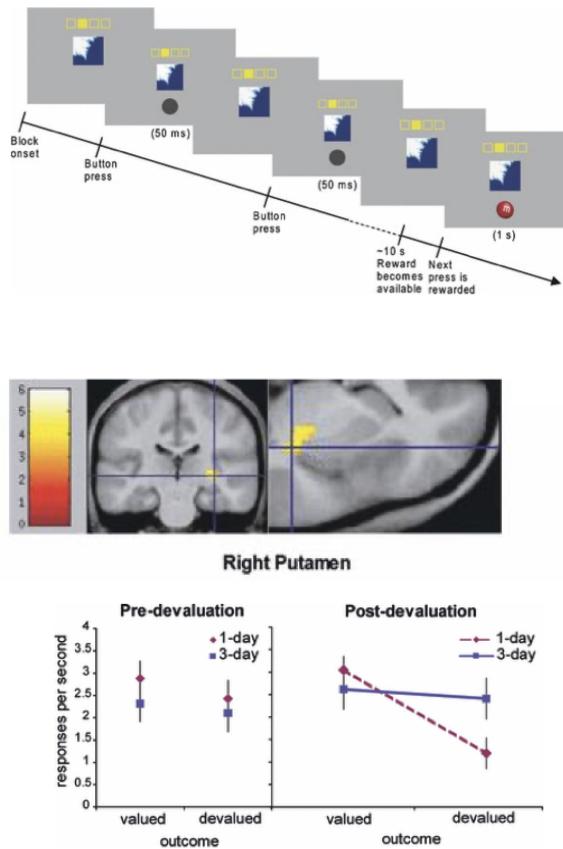


Behavior: how do the choices evolve? Subjects favor the high-probability reward choices. When you watch subjective pleasantness (how much do you like the common reward (orange juice), devalued (tomato), valued (chocolate), nothing) throughout the experiment, at the beginning, all rewards are above 0 and the no-opt was disliked (below 0). After devaluation, the devalued reward pleasantness changes drastically.

So the behavior was goal-directed, as expected.

Results at neural level, asking the question “Are there brain areas that respond differently between the still motivationally relevant outcome and the devalued one?” → Medial Orbito-Frontal Cortex - structure related to mesocortical pathways. Evidence for goal directed behavior.

Make an extensive training to see if the behavior gets habitual. Two groups, one of them trained 6 times more than the other one, using visual stimuli. Different stimuli cue suggested what action to perform to obtain food. The two outcomes were then devalued until you didn't want them anymore. Then put back to perform an extinction test.



Results (on responses per second): before the devaluation procedure, the two groups behave similarly, with no significant differences in response rates between groups or when responding for the two food rewards . After the devaluation procedure, the lightly trained group decreases the response for the devalued outcome, while the habitual group keeps it up. The still-valued outcome-related behavior is maintained by both groups.

Significant voxel clusters, including a region within the dorsolateral striatum (DLS) in the right posterior putamen-globus pallidus

Generation 3: Model-based vs Model-Free Computational Analyses

The third generation deals with the computational formalization of goal-directed (model-based), habitual (model-free) actions and their interactions.

A model-based algorithm selects actions by using a model to predict the consequences of possible courses of actions in terms of future states and the reward signals expected to arise from those states. A model-free algorithm selects action values for all the state action pairs obtained over many learning trials. When the environment of a model-based agent changes the way it reacts to the agent's actions, the agent can update the value (policy) of future states without the need to move to them. When the environment of a model-free agent changes the way it reacts to the agent's actions, the agent has to move to that state, act from it, possibly many times, and experience the consequences of its actions.

Latent Learning in Humans

The experimental task was a sequential two-choice Markov Decision Task in which all decision states are represented by fractal image. Design follows that of a binary decision tree. Each trial begins in the same state and subjects can choose between a L and R button press.

With a certain probability they reach one of two subsequent states in which they can choose again. Finally, they reach one of three outcome states associated with different monetary rewards. The experiment proceeded in two fMRI scanning sessions of 80 trials each.

- In the first session, subject choices were fixed and presented to them below the fractal image; however, subjects could still learn the transition probabilities.
- Between scanning sessions, subjects were presented with the reward schedule that maps the outcome states to the monetary payoffs. This mapping was rehearsed in a short choice task.
- Finally, in the second scanning session, subjects were free to choose left or right actions in each states and received the payoffs.

Results:

- Behavioral from free-choice session → Test if participants were able to make optimal choices by combining the knowledge they acquired about state transitions and reward contingencies. Any successful learning would be possible with model-based, but not model-free learning, as model-free learning focuses exclusively on predicting rewards, so it learns only if the rewards are given.
- Computational from free-choice session → Choice behavior during the entire session was best explained by hybrid model that integrates both Reward PE (model-free) and State PE (model-based).
- Neural → parameters estimated from computational models were used to find activations that correlated with SPE and RPE. Significant effects for SPE bilaterally in the intra-parietal sulcus and lateral prefrontal cortex, while significant effects for RPE in the ventral striatum

Sequential two-choice Markov Decision Tasks

They have been developed to discern the influence of model-free vs model-based controller on behavior and to determine whether neural signals are correlated with predictions and prediction errors specific to each controller. The usual task is “Maximize the reward”.

Subjects choose between two fractals, which probabilistically determined whether they would transition to the orange or blue second stage state. Action at the first state is associated with one likely and one unlikely transition. These probabilities were fixed and could be learned over time.

In the second stage state, subjects chose between two fractals, each of which was associated with a distinct probability of being rewarded with a 25 cents coin.

To incentivize subjects to continue learning throughout the task, the chances of pay off associated with the four second-stage options were changed slowly and independently.

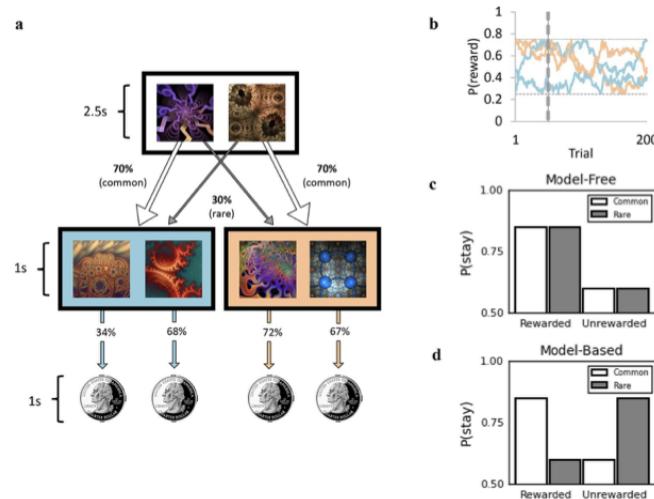
The chance of winning is almost stochastic.

Model-Free and model-based agents differ in the action selected after a rare transition:

- Model-Free agent ignores transition structure and prefers to repeat actions that lead to reward, irrespective of the likelihood of the first transition
- Model-Based agent respects transition structure and can ascribe rewards following a rare transition to an alternative (non-selected) action, which, despite not predicting reward on the current trial, will be more likely to lead to reward on future trials.

(c) Schematic representing the performance of a purely 'model-free' learner, who only exhibits sensitivity to whether or not the previous trial was rewarded vs. unrewarded, and does not modify their behavior in light of the transition that preceded reward.

(d) Schematic representing the performance of a purely 'model-based' learner, who is more likely to repeat an action (i.e. 'stay') following a rewarded trial, only if the transition was common. If the transition to that rewarded state was rare, they are more likely to switch on the next trial.



Results:

- Analysis of choice behavior → simple reinforcement predicts that a first-stage choice resulting in reward is more likely to be repeated on the subsequent trial, regardless of whether that reward occurred after a common or rare transition. Model-based prospective evaluation instead predicts that a rare transition should affect the value of the other first-stage option, leading to a predicted interaction between the factors of reward and transition probability.
- Computational → Choice behavior during test was best explained by a hybrid model that integrates both a RPE and SPE
- Neural → Parameters estimated from computational models were used to find activations that correlated with SPE and RPE. Activity in striatum occurred both for model-free and model-based prediction error. The activity correlated with the extent to which that subject's behavior was model-based.

Conclusions

Generation 3 results challenge the notion of a separate model-based vs model-free learner and suggest a more integrated computational and neural architecture for high-level human decision-making. 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.