# The shocking origin of neuroscience

If one were to set out in search for the earliest scientific breakthrough that led to this work an educated guess would land on Luigi Galvani. The legacy of the Italian polymath is grounded in his discovery that the muscles of frogs twitch when electrically stimulated (Galvani, 1791). He thereby refuted the contemporary belief that animal spirits inside hollow nerves drive movement and sensation. Although he was wrong in attributing the muscle twitches to an innate force he called “animal electricity”, he still managed to demonstrate the electrical nature of nerve impulses, thereby laying the foundation of electrophysiology and modern neuroscience (Piccolino, 1998).

The impact of Galvani's work was so immense that it has been likened to the French Revolution (Piccolino, 1998). It ultimately led Alessandro Volta to invent the electrical battery and inspired Mary Shelley to write the classic horror story *Frankenstein* (Piccolino, 1998). Galvani's name survives until today in the verb *galvanize* and still has a place in popular culture through songs such as Galvanize (curiously by The Chemical Brothers).

# A walk down memory lane

We experience the world around us filtered through the lens of our experiences. Without memories we could not hold on to these experiences locking us perpetually in the present. It follows that memories are at the core of what makes us humans.

Episodic memory is the ability to encode and later recollect experiences that contain a what, where and when. They are rich in detail, integrating information from multiple modalities, they are encoded automatically, require no repetitions, and can last an entire life (xx). By remembering these episodic memories, it is as if we were mentally transported back to that time, re-experiencing them anew.

An example for an episodic memory is when I was sitting in a small coffee shop in Sevilla in the company of my loved one. The sun had not yet reached its peak and was pleasantly warm. A mild breeze carried over the smell of freshly brewed coffee and bits of conversations from other patrons. It was a satisfying way to start the day, my body still exhilarated from the workout we just finished. The waiter brought over two coffees. On the way back to the kitchen he hesitated, turned on his heel and walked back to our table. "Your PhD thesis was a fantastic read", he said with a slight Spanish accent, adding "but why was your example for episodic memories so long?".

Semantic memories on the other hand refer to factual knowledge and understanding of concepts (such as knowing that the very real coffee shop in the above story was called "La Nueva Peseta"; xx). Together with episodic memories they belong to the subgroup of declarative memories (xx). ~~Sometimes declarative memories are termed explicit memory because they require explicit perception (?? xx).~~

Semantic and episodic memories are usually treated as different subcategories of declarative memories. In reality the line separating the two can get blurry. For example (xx).

Declarative memories can in turn be distinguished from non-declarative memories (xx). This category contains procedural memory (e.g., knowing how to make a coffee) and priming, which refers to the phenomena that exposure to a stimulus influences the behaviour or response to a later stimulus (e.g., judging someone’s character as "warmer" after holding a warm coffee; 10.1126/science.1162548). These memories do not require conscious perception which is why they are also referred to as implicit memories.

# Intracranial EEG (iEEG) and epilepsy

Roughly 1% of the population suffers from epilepsy, and in one-third of these cases treatment and medication provide no remedy from seizures [Kwan, 2011; n Engl J med]. If the seizure onset is focal, i.e., spatially confined it is sometimes possible to resect the epileptic tissue which effectively cures the patient [engel 1996, n engl j med j]. Henry Molaison, also known as Patient H.M., was the most renowned epilepsy patient. He underwent a resection of both hippocampi and large parts of his MTL (xx), which led to a seizure-free life. As a side effect of the surgery, he developed a graded retrograde amnesia and a complete anterograde amnesia inspiring a new wave of research implicating the hippocampus and neighbouring structures in episodic memory processing (xx). Nowadays, an extensive battery of tests is administered prior to resection, with the aim to exclude as much healthy tissue as possible (10.1038/s41593-018-0108-2). One important procedure is the transcranial implantation of depth electrodes at suspected seizure onset zones, based on seizure characteristics, anatomical scans, and long-term surface EEG recordings (xx). Once implanted these electrodes typically remain in place for 1-2 weeks to gain an understanding which brain regions are responsible for the generation of epileptic seizures and will later be resected.

While these electrodes are implanted, researchers perform experiments with willing patients granting insight into the neurophysiological underpinnings of various brain functions.

Ward and Thomas (1955) were the first to successfully record human single neurons. They did so in the posterior temporal lobe using glass micropipettes while surgeons tried to localize the epileptic focus and repair a bone defect in the patient’s skull. The type of microwire electrodes that are still in use today, (Fried et al., 1999) have been described in the early 70s by Babb and colleagues (Babb et al., 1973; electroenc & clinic neurophysiol). ~~In humans the most commonly used depth electrodes are of the~~ **~~Behnke-Fried type~~** ~~(see Figure xx) fabricated by Ad-Tech Medical Instrument Corporation.~~

These single-use intracranial depth electrodes consist of a 1.3 mm hollow macroelectrode through which a bundle of eight high-impedance microelectrodes and one low-impedance microwire is inserted. By default, the low-impedance wire is used as a reference for the high-impedance wires. Microwires have a width of 40 µm and radially protrude 4-5 mm past the end point of the macro depth electrode. They are made from platinum, which has a high impedance for lower frequencies and a low impedance for higher frequency bands. This allows the recordings of action potentials of multiple local single neurons superimposed on local field potentials.

Each microwire bundle typically yields around a dozen separate neurons. Usually, fewer single neurons can be recorded at the end of the first recording week, which is likely due to gliosis at the microwire tip (Fried et al., 1999).

Newer probes such as the Neuropixels 1.0 probe contains 384 channels across a 24 µm × 70 µm × 10 mm shank (xx), providing an advantage over conventionally used microwires which can spread out during electrode implantation in an unpredictable way (does anyone mention that? maybe 1999). As a consequence, spike detection and clustering cannot rely on local similarities between electrodes. Not only does this new probe allow for a higher quantity of recorded neurons, but the rigid distance between neighbouring channels (20 µm) allow a higher quality spike sorting as spikes are propagated across contacts.

Using a Neuropixels probe Durand and colleagues (Durand et al., 2022) recorded almost 600 neurons across 13 different brain regions using six different Neuropixels probes in a mouse.

In the first reported use of this novel probe in humans, Paulk and colleagues recorded upwards of 300 cortical single neurons in two patients awaiting DBS implantation for movement disorder. However, in one epilepsy patient awaiting tissue resection, the probe in the lateral temporal lobe only recorded the activity of 29 neurons (Paulk et al., 2022 Nature Neuroscience). Of note, the entire recording was conducted within the confines of the operating room for just 15 minutes, so no experimental intervention was possible (Paulk et al., 2022 Nature Neuroscience).

Compared to commonly used electrodes in humans, the higher yield of neurons with newer probes will facilitate analyses of assemblies of neurons and their interactions with different brain regions (Durand et al., 2022).

The clear advantage of intracranial electrophysiological recordings over traditionally used non-invasive methods is a spatially confined and well localized signal (vs. surface EEG or MEG) with a high temporal resolution (vs. fMRI) (10.1016/j.cell.2019.10.016). In contrast to invasive recordings in animals, humans can typically perform a task after minimal instructions and can provide comprehensible verbal feedback when prompted.

A severe disadvantage of intracranial recordings is a relatively limited coverage of the brain compared to traditionally used brain recording methods. This downside is exacerbated by the fact that the spatial positions of the intracranial electrodes are determined by clinical need and not scientific experimentation. Furthermore, access to epileptic patients that are willing to participate in scientific research is limited. Finally, even if these hurdles are overcome, it is important to ascertain that pathologic epileptic activity does not influence the obtained results 10.1016/j.cell.2019.10.016; Kastner paper).

# Microwire recording – LFP and Single Units

The recorded signal from the microwires can be divided into two components depending on their frequency. The first component is the local field potential (LFP), which reflects changes in the extracellular membrane potential and ranges until 300 Hz. Superimposed onto the LFP is the activity of individual neurons and multiunits in close proximity to the microwire.

Action potentials (also called *spikes*) are characterized by a steep and transient amplitude increase in the signal. Spike detection and sorting can be implemented using a variety of existing toolboxes, with new ones being developed continuously that demonstrate promising results (10.1101/2023.01.07.523036). Here, we used the wave\_clus algorithm, which is described in detail in Chaure, Rey and Quian Quiroga (2018). The following is a brief synopsis of the processing steps performed by this algorithm. The first step to detect these neural spikes is to filter the data so it only contains the spike-band which ranges from 300 Hz to 3000 Hz. Next, the data is segmented into smaller epochs of typically five minutes each, so artefacts occurring in one segment do not increase the threshold across the entire recording. Each one of these epochs is then individually thresholded using some form of deviation to a measure of central tendency (such as the mean or median). Points where the threshold is surpassed are stored as putative spikes. This spike detection is done separately for positive and negative deflections. Once a spike is identified, 64 data points around the maximum are extracted, which corresponds to a 2 ms window at a sampling rate of 32000 Hz. The spike peak is aligned to the 20th sampling point. To avoid misalignment of the spike the waveshape is first upsampled to 320 data points using cubic spline-interpolated waveforms and then downsampled again. Based on the extracted spike-waveform, features are computed using a four-scale multiresolution decomposition with a Haar wavelet. This results in 64-wavelet coefficients for each spike. The 10 most significant coefficients are identified using a Lilliefors test and used for the clustering procedure (Chaure et al., 2018). Nonparametric clustering in the feature space is done using superparamagnetic clustering (SPC). SPC groups spike waves into clusters based on nearest-neighbour interactions (Blatt et al., 1996). Template-matching in Euclidian space is performed to assign unclassified waveforms to one of the identified clusters. The resulting clustering solution is then manually inspected and further optimized by rejecting artefact cluster, splitting clusters that represent multi-unit activity and merging clusters that likely stem from the same neural source.

The extracellularly recorded local field potential (LFP) represents synchronously active neurons that are spatially aligned. Synaptic activity is the largest contributor to the LFP, but transmembrane currents from soma, dendrites, spikes, and spike afterpotentials also impact the LFP (xx). The LFP can be divided into periodic (oscillatory) and aperiodic (fractal, non-oscillatory) components. Aperiodic power is inversely related to the frequency and roughly follows a 1/f relationship (where f is the temporal frequency). This power-frequency relationship is likely due to dendrites acting as a low-pass filter (10.1007/s10827-010-0245-4; re-read abstract) and because fewer neurons can be active in shorter cycle lengths of higher frequencies (10.1038/nrn3241).

In the past, the aperiodic part of the signal was often ignored or considered background noise (xx). However, more recent research has pointed to the steepness or tilt as well as the offset of the 1/f aperiodic component as an indicator for excitation (xx) and a proxy for neural firing (Manning xx).

The periodic part reflects true oscillatory activity (i.e., rhythmic activity in a circumscribed frequency range). Activity in these narrowband frequencies have been associated with a wide range of cognitive processes (xx) and states (xx). Analysing this oscillatory activity without consideration of the 1/f shape can be problematic (10.1016/j.neuroimage.2022.118929; Herwig paper xx) as the shape of the 1/f can bias the oscillatory activity. Moreover, a tilt or change in offset can be erroneously interpreted as a change in oscillatory activity (Herwig paper xx).

There are multiple methods to separate the signal into periodic and aperiodic parts such as Irregular Resampling Auto-Spectral Analysis (IRASA) (10.1371/journal.pone.0024331) and Fitting Oscillations and One Over F (FOOOF; 0.1038/s41593-020-00744-x).

# The hippocampus

The etymological root of *hippocampus* comes from the Greek words "hippos" (horse) and "kampos" (sea monster) and trace back to the anatomist Julius Caesar Aranzi, who compared the shape of the hippocampus to that of a sea horse (10.3171/2014.11.JNS132402). Although the term hippocampus prevailed, different names have been proposed in the past, such as “silkworm” or “Ram’s horn” (10.3171/2014.11.JNS132402).

Humans have two mirrored hippocampi, one in each hemisphere. These hippocampi are located beneath the neocortex within the medial temporal lobe (MTL).

The hippocampus can be divided into the dentate gyrus, hippocampus proper (CA1-CA3) and the subiculum. Highly processed information flows from prefrontal neocortex, perihinal cortical areas and association cortices through the EC to the hippocampus (Teyler & Rudy 2007). This cortical information is integrated with subcortical input from the amygdala and thalamus (Teyler & Rudy 2007). This positions the hippocampus uniquely to integrate cortical and subcortical information streams (Swanson & Mogenson, 1981 xx).

# The Indexing Theory of the human hippocampus

More than three decades ago, Teyler and DiScenna proposed the Indexing Theory as a framework to explain hippocampal function during the encoding and retrieval of episodic (at the time called experiental) memory in humans (Teyler & DiScenna, 1989).

According to the Indexing Theory, during initial encoding the various multimodal elements that make up an episode instate a cortical activity pattern that is projected to an assembly of neurons in the hippocampus.

Subsequently, a partial input of the initial experience is sufficient to reactivate the entire assembly of associated hippocampal neurons, a process known as pattern completion. These neurons then project back to the neocortex, reinstating the entire experience.

Pattern separation refers to the complementary ability to distinguish between distinct episodes. Because each experience is uniquely indexed even the highly overlapping cortical representations of two similar episodes can be separated in the hippocampus.

This hippocampal index allows a flexible way to quickly store the cortical representation of an episodic memory. Over time, the initially strengthened synaptic connections for unimportant memories either decay (xx) or fall victim to interference (xx).

Within this framework the function of the hippocampus can be likened to a librarian: it can direct one to the necessary information within the library (the neocortex) but does not possess the knowledge itself. This implies that the hippocampus is content-free and does not contain semantic information.

# Information flow during memory processing

According to a model by O’Reilly and Rudy (2001), during memory encoding information from the cortex reaches the entorhinal cortex (EC) where two representations are generated. One representation is projected via the broad and diffuse perforant path to the dentate gyrus (DG), forming a sparse rendering of the cortical activity pattern. The DG then connects to CA3 through the sparse, focused and topographically arranged mossy fibre pathway, with approximately 70 synapses linking to each CA3 neuron in rats (xx). At the same time, the other representation invertibly projects from the EC to CA1. This connection is point-to-point and not diffuse like the perforant path (Tamamaki, 1991 xx). Due to the coactivity of neurons in CA3 and CA1, their diffuse and widespread synaptic connections through the Schaffer Collaterals are strengthened.

During retrieval, a partial input of the original representation is sufficient to reactivate the representation in CA3, where the entire representation is pattern completed. This in turn reinstates the appropriate CA1 representation that can project back to the EC because of the bidirectional connection between EC and CA1.

# Different types of memory

The Standard Model of Systems Consolidation (Squire & Alvarez, 1995) proposes that a memory trace is initially encoded in the hippocampus and only weakly encoded in the cortex. Over time the hippocampus reactivates the cortical pattern thereby gradually strengthening the synaptic connections that formed the initial memory trace in the cortex - a concept that dates back to Marr (1971). As a result, the hippocampus eventually becomes redundant. This is in line with the graded retrograde amnesia observed in patient H.M. (Scoville & Milner, 1957), whose hippocampus and extensive parts of the medial temporal lobe (MTL) had been removed. The forgetting of more recent memories can be attributed to their incomplete consolidation.

McClelland and colleagues (McClelland et al., 1995) extended the Standard Model of Systems Consolidation and developed a computational theory wherein the hippocampus is responsible for rapid learning of new information that could then be integrated in the neocortex over longer time periods. This would prevent catastrophic interference between older and newer memories in the cortex. The hippocampus separates experiences and avoids interference through the implementation of a sparse and orthogonal code where each event is represented by a distinct assembly of neurons. The neocortex and hippocampus as Complementary Learning Systems provide a solution to the challenge that the brain needs to both recognize general patterns in the environment and capture the details of a particular episode (O’Reilly & Rudy 2001).

In contrast, the Multiple Trace Theory (Nadel & Mocovitch, 1997) proposed that the hippocampus remains essential for episodic memory even for remote memories. However, similar to the Systems Consolidation account, the hippocampus aids in the stabilisation of semantic memories in the neocortex. In support of this Corkin (2002) argued that remote memories of H.M. were semanticized and thus did not reflect retrieval of true episodic memories.

Importantly, whether the Systems Consolidation or the Multiple Trace Theory prevails has no bearing on the concept of a hippocampal index assembly which is compatible with either framework.

# How are neurons allocated to a memory trace?

Over one hundred years ago Richard Semon proposed that a memory is represented by the long lasting physical changes in neural assemblies that encoded the initial experience. This memory trace is termed “engram” in the animal literature (Richard Semon, 1921; doi.org/10.1038/npp.2016.73 xx, doi.org/10.1038/nrn4000 xx).

Unlike Index Neurons, which are assumed to be in the hippocampus, the entire engram representing an experience spans multiple assemblies in various brain regions that are functionally connected (Roy and colleagues doi: 10.1101/668483).

Optogenetics and chemogenetics (used in humans? xx), which are not available in human research, have been especially beneficial to memory research in animals. Although findings from rodent brains do not by default translate to the human brain, there is enough overlap that (non-human) animal work can inform human research.

Experiments conducted on rodents revealed that neurons are allocated to an engram based on their excitability, with those having higher excitability more likely to be included (10.1038/npp.2014.234, doi.org/10.1503/jpn.100015 xx). Excitability is defined as the inclination of a neuron to fire an action potential in response to a signal (Dong et al. 2006). Rashid and colleagues (Rashid et al., 2016) showed that neurons assigned to an engram inhibit neighbouring neurons for about 6 hours through GABAergic interneurons. Without this inhibition, memories that occur close in time might be encoded by non-overlapping neurons.

After being allocated to an engram, neurons representing an event remain in a state of elevated excitability for over six hours. Consequently, some of the initial engram neurons are likely to be coallocated to events that occur within this timeframe (Cai et al., 2016; Rashid et al., 2016). After this period, excitability drops making it less likely that the same engram neurons represent temporally distant events (Frankland & Josselyn 2015; Silva et al., 2009).

Cai and colleagues (Cai et al., 2016) found evidence for this in CA1 of mice, that were presented with context A, followed by context B seven days later and then context C five hours later. Engrams representing the contexts separated by a shorter temporal gap were largely overlapping, while those with a larger time delay showed no such overlap. Rashid and colleagues (xx) extended these findings by optogenetically stimulating neurons in the lateral nucleus of the amygdala that were allocated to an event 24h before a second event took place (i.e., outside of the 6 hour window of increased excitability). Due to this artificially induced excitability the second event was coallocated to the same subset of neurons. A similar result was obtained when the remote memory was retrieved prior to acquisition of a related memory, suggesting a mechanism for integrating newer memories with relevant older memories (Rashid et al., 2016; Yokose et al., 2017: two distant memories show an overlap if they are co-retrieved).

This mechanism of coallocation is suspected to be responsible for false memories: engram cells in the dentate gyrus active during the exploration of context A were optogenetically reactivated in context B, where the mice also received footshocks. Mice then showed fear reinstatement in context A (artificial fear memory) and B (natural fear memory), but not in a third neutral context (10.1126/science.1239073). Similarly, Vetere and colleagues (10.1038/s41593-019-0389-0) tagged neurons in the olfactory bulb and synchronized it with either appetitive or aversive neural pathways. Subsequently mice showed attraction or aversion to the real odour giving credence to the idea that an artificial memory was created the absence of a real experience.

Engram neurons are necessary and sufficient for memory retrieval. After destroying a subset of neurons that were initially allocated to a fear memory mice suffered from a profound memory loss (Han et al., 2009 10.1126/science.1164139;). Importantly this loss-of-function was specific to the fear memory and new fear conditioning was possible. Ablating other neurons did not lead to a disruption in memory. Conversely, artificial reactivation of engram cells in the dentate gyrus reliably led to the retrieval of the memory even in the absence of external retrieval cues (10.1038/nature11028). In a neutral context mice did not freeze until the engram representing the fear memory was optogenetically reactivated. This represents a gain-of-function and cements engram cells as causally relevant for memory processing.

# Neurons coding content: Concept Cells

Concept cells are neurons in the human MTL that fire in response to specific concepts in an all-or-none way (Rey et al., 2018). They exhibit a high degree of multimodal invariance (i.e., they respond to Jennifer Aniston as an image or her spoken name) and context invariance (i.e., a concept neuron tuned to Jennifer Aniston would activate regardless of whether you see her in a movie, a park or in a café) (xx).

Curiously, the latency of their firing rate is much later than would be required by simple sensory processing and object recognition, which is an indication of their involvement in memory processing (Mormann et al., 2008). This lines up with the observation that most concept neurons are tuned to personally relevant concepts and depend on the subjective and conscious perception rather than objective sensory properties (Quiroga et al., 2014, 2008).

These concept neurons are not topographically organized, i.e., spatially close concept neurons might code for vastly different concepts (Quiroga 2016?). This spatial organization benefits episodic memory processing as it allows association between any two concepts without connecting distant areas (plugging into memory xx). According to Quian Quiroga (Cell 2019; tics; 2012 paper) these CN are the building blocks of episodic memory formation and retrieval. If you met your best friend in your favourite café the concurrent activation of two assemblies of CN (one for your friend and one for the café) would represent the episode in the hippocampus. These assemblies would then project back to the neocortex reinstating the sensory activity pattern first induced during the formation of the episode. This back-projection parallels the one described in the Indexing Theory (Teyler 1 & 2 xx) with the important difference that the hippocampal representation consists of previously existing concept specific representations/assemblies.

A separate memory of the same friend in a park would in turn be represented by the simultaneous activity of the same assembly coding for your friend and another assembly representing the park.

# Additional notes:

Subfields and how they are connected add citation:(doi.org/10.1016/j.cub.2015.10.049).

CA3 has a vast autoassociative network with synapses onto other CA3 neurons. This physioologiy inspired many impoortant bla bla (already in marr)

Medial septum makes theta in hippocampus

Research into different kind of neurons has led to a veritable embarasse de richesse

Episodic memories although originally defined as such, can implicitly considered conjunctive codes (Oreily & rudy paper). Complementary learning systems: O’Keefe & Nadel (1978): taxon local | hirsh 1974 | McClelland et al. (1995)

The hippocampus automatically binds sensory elements into a conjunctive code which corresponds well with the definition of episodic memory (O’Reilly & Rudy 2001).

In 1929 Hans Berger published his seminal work where he recorded electric potentials on the human scalp using an electroencephalopgraph (german: Elektroenkephalogramm; Berger, 1929). He mostly observed oscillations around 10 Hz, which he therefore termed alpha oscillations.