The hippocampus as an indexing machine of episodic memory

Luca Dominik Kolibius

A thesis submitted to the University of Glasgow for the degree of Doctor of Philosophy (PhD)

School of Psychology and Neuroscience
College of Medical, Veterinary & Life Sciences
University of Glasgow

February 2023

Abstract

Episodic memories refer to our ability to encode and reinstate experiences. By extension, these memories shape how we view ourselves and the world around us. Despite this, little is known about how neurons in the hippocampus encode and retrieve new episodes. Here, I will demonstrate evidence for single neurons in the human hippocampus that code specific episodic memories (hence called Episode Specific Neurons), both through a rate code and a temporal code. Importantly, these neurons cannot be construed as coding for specific time-points or concepts. Next, I will extend these findings to population activity in the local field potential. I report evidence for a reinstatement in high frequency power during successful memory processing that mirrors earlier findings in single neurons. Again, these results cannot be explained by activity induced by a content-code. Despite the undisputed importance of theta activity in memory processing, we find no consistent evidence of an increase in theta power during memory processing. Likewise, we find no evidence that earlier identified Episode Specific Neurons or other hippocampal neurons fire preferentially at a particular theta phases or theta phase offsets between encoding or retrieval of episodic memories. Lastly, I embed these findings in the broader literature, identify future experiments, and discuss possible translational applications.

Acknowledgements

This manuscript is the culmination of over four years of work. Naturally, there are many people that supported me over these years and who I wish to thank.

First, I would like to thank my wife María who has been an incredible support over the years. Through the dark Scottish winter months, you have been my ray of sunlight. I do not know a single person as strong and ambitious as you and I am grateful to have you by my side. I love you.

I thank my family, especially my parents Verena and Michael Kolibius for their unwavering support over my entire life and for being a permanent reference point in an otherwise chaotic world. I could not have done this without you. I know that you will always be there when I need you. One day I hope to be such a good parent to my children.

At the very beginning of my PhD Simon Hanslmayr and Bernhard Staresina invited me to a meeting discussing which dataset should be used to investigate “Index Neurons”. I was unrecoverably lost which found its peak when Bernhard understood some implications before Simon finished his sentence. I doubted myself. I soon recovered my excitement for research but often stumbled on unseen ground.

I am very grateful to Simon, my *Doktorvater* for his guidance and his trust in me even when I was hard stuck on a particular problem for weeks or head over heels down a rabbit hole. For me, Simon embodies coolness and a keen mind in equal parts. You have been a fantastic and supportive supervisor. I would also like to thank my second supervisor Howard Bowman. I find your intuitive grasp on mathematics inspiring and thoroughly enjoy our long conversations about neuroscience and all other things.

I thank the sharp minds that have accompanied me in this journey and had so many valuable conversations, especially Sander van Bree, Mircea van der Plas, Marit Petzka, Chris Postzich. Thank you for listening to my ecstatic outbursts on ideas that are often little more than conjecture (but sometimes turn out to be useful!).

Finally, I wish to thank all the patients who have participated in our studies. Without you, this work would have never been possible.

Declaration

All work in this thesis was carried out by the author unless otherwise stated. This work has not been submitted for any other degree.

List of Abbreviations

BL	Baseline
CN	Concept Neuron
CSW	Concept Specific Microwire
EC	Entorhinal Cortex
EEG	Electroencephalogram
Enc	Memory Encoding
ep	Episode
ESN	Episode Specific Neuron
ESW	Episode Specific Microwire
fMRI	Functional magnetic resonance imaging
FOOOF	Fitting Oscillations and One over F
HFP	High Frequency Power
iEEG	Intracranial Electroencephalogram
IRASA	Irregular Resampling Auto-Spectral analysis
LFP	Local Field Potential
MTL	Medial Temporal Lobe
nESN	ESN Non-Reinstated Episodes
PhD	Doctor of Philosophy
PTSD	Posttraumatic Stress Disorder
rESN	ESN Reinstated Episodes
Ret	Memory Retrieval
SPC	Superparamagnetic Clustering
SU	Non-Tuned Single Unit (Non-ESN)
TC	Time Cell
tESN	temporal Episode Specific Neuron

Table of Contents

List of Tables

List of Figures

Chapter 1

General Introduction

1.1 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).

An equally reasonable choice would be Ramón y Cajal, who in the late 1880s discovered that the central nervous system is made up of separate interconnected neurons (Ramón y Cajal, 1888). He later suggested that their ability to form new connections may explain learning (Ramón y Cajal, 1894). About 60 years later, Donald Hebb proposed a principle according to which the connections of two neurons are strengthened if one neuron repeatedly excites the other neuron (Hebb, 1949). This led to the popular paraphrase "Neurons that fire together, wire together" which was coined not by Hebb himself, but by Shatz (Shatz, 1992).

1.2 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, a term coined by Endel Tulving in 1972 (Tulving, 1972), 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 (Teyler & Rudy, 2007; Tulving, 2002). By remembering these episodic memories, it is as if we were mentally transported back to that time, re-experiencing them anew (Tulving, 2002). An example of 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"; Squire, 1987). Together with episodic memories they belong to the subgroup of declarative memories (Squire, 1987). Declarative memories, sometimes also called explicit memories, can be expressed (i.e., declared) overtly and form the basis for conscious recollection (Squire, 1987, 1992). In reality, the line separating semantic and episodic memories can get blurry. For example, if you were asked how old you were when you received your childhood pet, the retrieved memory would have semantic (your age) and episodic aspects (the experience itself).

Declarative memories can in turn be distinguished from non-declarative memories (Squire, 1992). 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; Williams & Bargh, 2008). These memories do not require conscious perception which is why they are also referred to as implicit memories (Squire, 1992).

1.3 Intracranial electrophysiological recordings in epilepsy patients

Roughly 1% of the population suffers from epilepsy, and in one-third of these cases treatment and medication provide no remedy from seizures (Kwan et al., 2011). 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). Henry Molaison, also known as Patient H.M., was the most prominent epilepsy patient. He underwent a resection of both hippocampi and large parts of his MTL, which led to a seizure-free life (Corkin, 1984; Scoville & Milner, 1957; Squire, 2009). As a side effect of the surgery, he developed a graded retrograde amnesia and a complete anterograde amnesia, meaning that he retained some distant memories, but could neither remember recent memories nor create new ones (Corkin, 1984; Scoville & Milner, 1957; Squire, 2009). This inspired a new wave of research implicating the hippocampus and neighbouring structures in episodic memory processing (Corkin, 1984; Scoville & Milner, 1957; Squire, 2009). Nowadays, an extensive battery of tests is administered prior to resection, with the aim to exclude as much healthy tissue as possible (Parvizi & Kastner, 2018). One important procedure is the intracranial implantation of depth electrodes at suspected seizure onset zones, based on seizure characteristics, anatomical scans, and long-term surface EEG recordings (Parvizi & Kastner, 2018). 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 (Parvizi & Kastner, 2018; Quian Quiroga, 2019). While these electrodes are implanted, researchers perform experiments with willing patients granting insight into the neurophysiological underpinnings of various brain functions.

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) (Quian Quiroga, 2019). 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 (Parvizi & Kastner, 2018; Quian Quiroga, 2019). 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 (Parvizi & Kastner, 2018; Quian Quiroga, 2019).

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). These so-called Behnke-Fried electrodes are single-use intracranial depth electrodes that 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 contain 384 channels across a 20 µm × 70 µm × 10 mm shank (Dutta et al., 2019; Jun et al., 2017). Apart from a higher quantity of recorded neurons, the rigid distance (20 µm) between neighbouring channels allows for a higher quality spike sorting as spikes are propagated across contacts. In comparison, local similarities between microwires cannot be used in conventionally used electrodes as they spread out in an unpredictable way during implantation. 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). 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). 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).

1.4 Microwire recording – LFP and Single Units

The recorded microwire signal 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 multi-units 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 (Pachitariu et al., 2023). Here, we used the wave_clus algorithm, which is described in detail in Chaure and colleagues (Chaure et al., 2018). The following is a brief synopsis of the processing steps performed by this algorithm. The first step to detect 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 detected, the features of each spike-waveform are computed using a Haar wavelet and the most significant coefficients are identified using a Lilliefors test (Chaure et al., 2018). Next, nonparametric clustering is performed in the feature space using superparamagnetic clustering. Superparamagnetic clustering groups spike waves into clusters based on nearest-neighbour interactions (Blatt et al., 1996). Through template-matching in Euclidian space, unclassified waveforms are assigned to one of the identified clusters. The resulting clustering solution is then manually inspected and further optimized by rejecting artefact clusters, splitting clusters that represent multi-unit activity and merging clusters that likely stem from the same neural source (Chaure et al., 2018).

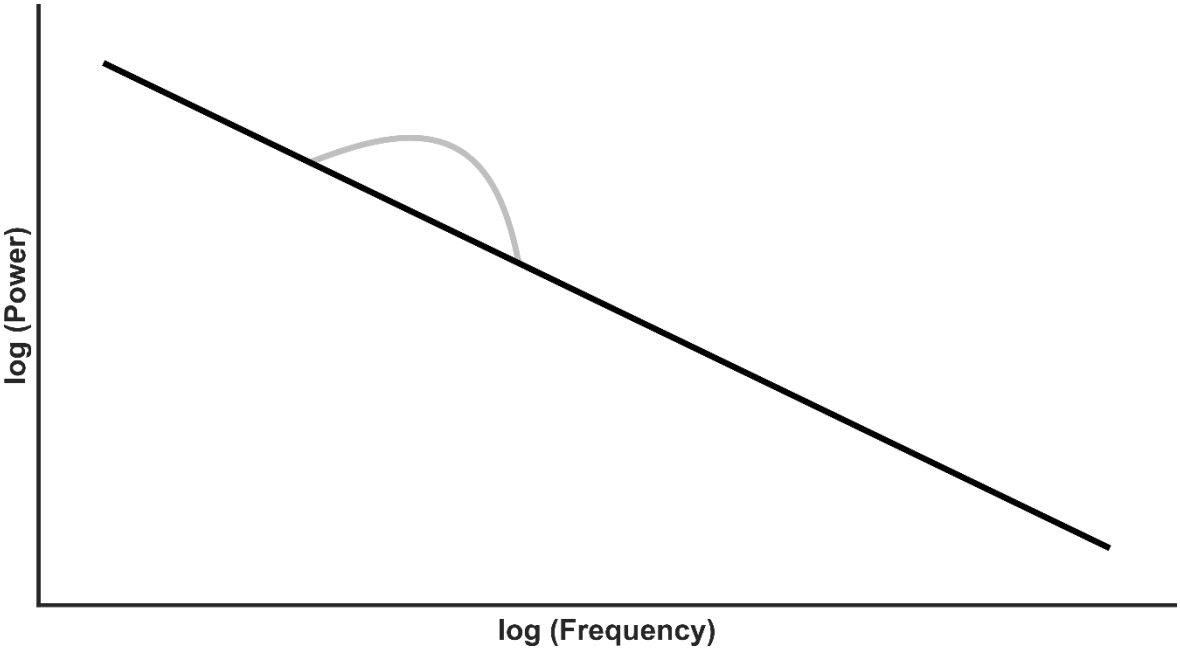


Figure 1.1. Example schematic of a power spectrum visualising the 1/f relation between power and frequency.
The black line represents the aperiodic component of the signal, and the grey line reflects the superimposed periodic activity. The x-axis shows the frequency, and the y-axis displays the power. Both axes are in log-space.

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 (Buzsáki et al., 2012). The LFP can be divided into periodic (oscillatory) and aperiodic (fractal, non-oscillatory) components (see Figure 1.1.). 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 (Linden et al., 2010) and because fewer neurons can be active in shorter cycle lengths of higher frequencies (Buzsáki et al., 2012). In the past, the aperiodic part of the signal was often ignored or considered background noise (Donoghue et al., 2020). 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 (Gao et al., 2017) and a proxy for neural firing (Manning et al., 2009). The periodic part reflects true oscillatory activity (i.e., rhythmic activity in a circumscribed frequency range). Activity in these narrowband frequencies has been associated with a wide range of cognitive processes (Buzsaki & Tingley, 2018; Hanslmayr et al., 2007; Jensen et al., 2007; Klimesch et al., 2007; Landau & Fries, 2012). Analysing this oscillatory activity without consideration of the 1/f shape can be problematic (Herweg et al., 2020; Samaha & Cohen, 2022) 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 (Herweg et al., 2020). There are multiple methods to separate the signal into periodic and aperiodic parts such as Irregular Resampling Auto-Spectral Analysis (IRASA; Gao et al., 2011) and Fitting Oscillations and One Over F (FOOOF; Donoghue et al., 2020).

1.5 The hippocampus

1.5.1 Etymology and anatomy

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 (Bir et al., 2015). Although the term hippocampus prevailed, different names have been proposed in the past, such as “silkworm” or “Ram’s horn” (Bir et al., 2015). 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 the prefrontal neocortex, perirhinal 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 (Swanson & Mogenson, 1981; Teyler & Rudy, 2007).

1.5.2 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 experiential) memory in humans (Teyler & DiScenna, 1986). 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 (Teyler & DiScenna, 1986; Teyler & Rudy, 2007). 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 (Teyler & DiScenna, 1986; Teyler & Rudy, 2007). 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 (Hardt et al., 2013) or fall victim to interference (Underwood, 1957). 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 as hippocampal neurons arbitrarily bind concurrent cortical activity irrespective of the semantic content they represent.

1.5.3 Information flow during memory processing

According to a model by O'Reilly and Rudy (O'Reilly & Rudy, 2001), during memory encoding information from the neocortex 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 (O'Reilly & Rudy, 2001). At the same time, the other representation projects from the EC to CA1 and back. This connection is point-to-point and not diffuse like the perforant path (Naber et al., 2001; Tamamaki & Nojyo, 1995). 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 (O'Reilly & Rudy, 2001).

1.5.4 Role in remote memories

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 neocortex. Over time the hippocampus reactivates the cortical pattern thereby gradually strengthening the synaptic connections that formed the initial memory trace in the neocortex – a concept that dates back to Marr (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 into the neocortex over longer time periods (see also Hirsh, 1974). The hippocampus separates experiences and avoids catastrophic interference between older and newer memories through the implementation of a sparse and orthogonal code where each event is represented by a distinct assembly of neurons. This complementary learning systems approach provides 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). Of note, Schapiro and colleagues (Schapiro et al., 2017) solved this conundrum in a computational model by assigning statistical learning and encoding of specific episodes to separate hippocampal subfields. In contrast, the Multiple Trace Theory (Nadel & Moscovitch, 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 (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.

1.6 Neurons in the hippocampus

1.6.1 Neurons coding content: Concept Neurons

Concept Neurons (CNs) are neurons in the human MTL that fire in response to specific concepts in an all-or-none way (Gelbard-Sagiv et al., 2008; Quian Quiroga et al., 2008; Quian Quiroga et al., 2005) 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; Quian Quiroga et al., 2005). 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 (Quian Quiroga et al., 2014; Quian Quiroga et al., 2008). These concept neurons are not topographically organized, i.e., spatially close concept

neurons might code for vastly different concepts (De Falco et al., 2016). This spatial organization benefits episodic memory processing as it allows association between any two concepts without connecting distant areas (Quian Quiroga, 2019). According to Quian Quiroga (Quian Quiroga, 2019, 2020; Quiroga, 2012), 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 & DiScenna, 1986; Teyler & Rudy, 2007) with the important difference that the hippocampal representation consists of previously existing concept-specific neural 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.

1.6.2 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 (Semon, 1904). This memory trace is termed “engram” in the animal literature (Josselyn et al., 2015; Park et al., 2016; Semon, 1904). 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 et al., 2019). Optogenetics and chemogenetics have been especially beneficial to memory research in animals. 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 (Frankland & Josselyn, 2015; Josselyn, 2010). 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 six 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 (Rashid et al., 2016) 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 six 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 the acquisition of a related memory, suggesting a mechanism for integrating newer memories with relevant older memories (Rashid et al., 2016; Yokose et al., 2017). 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 (Ramirez et al., 2013). Similarly, Vetere and colleagues (Vetere et al., 2019) 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 in 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). 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 (Liu et al., 2012). 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. 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 and provide useful hypotheses. For instance, it is unknown how neurons become assigned to a memory in humans, but it is possible that excitability determines this allocation process as well.

1.7 Goal of this thesis

Despite numerous studies highlighting the importance of the hippocampus for memory processing, the underlying coding mechanism is still elusive. The Indexing Theory of Teyler & DiScenna (Teyler & DiScenna, 1986) suggests the existence of neurons that are active during the initial encoding of a memory and later reinstated during the successful retrieval of the same memory. Although proposed over 35 years ago, no evidence for such a neural pattern has been found in humans.

In the second chapter of this thesis, I present evidence for Episode Specific Neurons (ESNs) that code for discrete episodes through a reinstatement of their firing rate from encoding to retrieval. Importantly, using various approaches I demonstrate that ESNs are different from Concept Neurons in that they code for a conjunctive representation of the elements within an episode, rather than a particular element or concept within that episode.

In the third chapter, I extend these single neuron findings to the population activity reflected in the local field potential (LFP). Power in the high frequency band (HFP; 40-200 Hz) has been used as a proxy for neural activity of the local field potential (LFP; Buzsáki et al., 2012). In line with the findings from Chapter 2, I identified a significant number of microwires that show a reinstatement of HFP from encoding to retrieval in individual episodes. Again, this code was not driven by content-specific activity.

In the fourth chapter, I report findings on the role of the theta band in memory processing. Recent research suggested that successful memory processing is reflected in an aperiodic power shift from the lower frequencies to the higher frequencies and a periodic narrow-band theta power increase. According to an influential theory memory encoding and retrieval occur in opposing phases of the theta oscillation to avoid catastrophic interference between new and older memories (Hasselmo et al., 2002). However, I found no conclusive evidence for a difference in aperiodic power or theta power between (i) remembered and forgotten episodes or (ii) between episodes that contained reinstating ESNs and episodes without ESN activity. Likewise, I did not discover consistent evidence of a (iii) phase preference of single neuron or ESN activity during encoding or during retrieval, (ii) or a significant theta phase offset between neurons firing at encoding and retrieval.

Finally, in the fifth chapter, I will summarize the findings of the preceding three chapters and propose potential avenues for future research. This includes possible experiments that may determine if ESNs satisfy the criteria for Index Neurons introduced by Teyler and Discenna (Teyler & DiScenna, 1986). Moreover, I will suggest a way how CN might originate from ESNs and outline ways in which the insight gained from the basic research in this thesis can be applied to aid patients suffering from dementia and PTSD.

Chapter 2

Hippocampal neurons code individual episodic memories in humans

Abstract

The hippocampus is an essential hub for episodic memory processing. However, how human hippocampal single neurons code multi-element associations remains unknown. Some argue that each hippocampal neuron codes for an invariant element within an episode. Instead, others have proposed that hippocampal neurons bind together all elements present in a discrete episodic memory. Here, we provide evidence for the latter. We show that individual neurons, which we term Episode Specific Neurons (ESNs), code discrete memory episodes. These ESNs do not reflect the coding of a particular element in the episode (i.e., concept or time). Instead, they code for the conjunction of the different elements that make up the episode.

2.1 Introduction

Episodic memory refers to our ability to reinstate the what, where and when of past experiences (Tulving, 2002). This ability is thought to depend on the reinstatement of neural activity that was present at memory encoding (Pacheco Estefan et al., 2019). It is undisputed that the hippocampus plays an integral role in episodic memory processing (Lisman et al., 2017; Marr, 1971; Squire, 1992) and the binding of multimodal information (Cooper & Ritchey, 2020). However, how it codes episodic memories remains controversial.

One important open question is whether neurons in the hippocampus code for specific elements or an entire episode. Concept Neurons in the hippocampus fire in response to specific invariant elements independent of the context in which they are presented (Gelbard-Sagiv et al., 2008; Mormann et al., 2011; Mormann et al., 2008; Quian Quiroga et al., 2005). One contemporary idea is that the diverse elements that make up an episode are coded by the simultaneous activity of a set of these Concept Neurons (Quian Quiroga, 2020; Quiroga, 2012) or by expanding the selectivity of existing Concept Neurons (Ison et al., 2015). According to this framework when you are sitting in your favourite coffee shop with your best friend, one set of Concept Neurons might code for the coffee shop and a separate set for your friend (Figure 2.1A).

Alternatively, single units in the hippocampus might sparsely encode a specific set of elements within an individual episode and act as pointers to cortical modules during memory reinstatement. According to this so-called Indexing Theory (Teyler and DiScenna, 1986; Teyler and Rudy, 2007), the entire episode with your friend in the coffee shop is represented by a set of hippocampal neurons (Figure 2.1A). Unlike Concept Neurons, these Episode Specific Neurons (ESNs) would fire in response to the conjunction of all the diverse information within an episode and not in response to individual content elements. Despite computational models pointing towards the existence of ESNs (Bowman & Wyble, 2007; Krotov & Hopfield, 2020; Parish et al., 2021; Whittington et al., 2022), to this day there is no evidence for such a sparse conjunctive code in humans.

In the present work, we provide support for the existence of this content-agnostic episodic memory code implemented through Episode Specific Neurons. We leveraged intracranial microwire recordings to investigate the firing patterns of neurons in the human hippocampus and hypothesized that a significant number of hippocampal neurons reinstate their firing rate within a specific episode (i.e., fire during encoding and retrieval).

Importantly, these ESNs would code for the conjunctive elements present within an episode and are not tuned to individual elements within the episode. The existence of ESNs does not

preclude Concept Neurons from participating in episodic memory processing. However, investigating the role of Concept Neurons in episodic memories goes beyond the scope of this work. As control analyses, we investigated whether this firing activity can be explained by a firing response to specific invariant elements, as occurs in Concept Neurons (Quian Quiroga et al., 2005), or by a time preference, as occurs in Time Cells (TC; Reddy, Zoefel et al., 2021; Umbach et al., 2020).

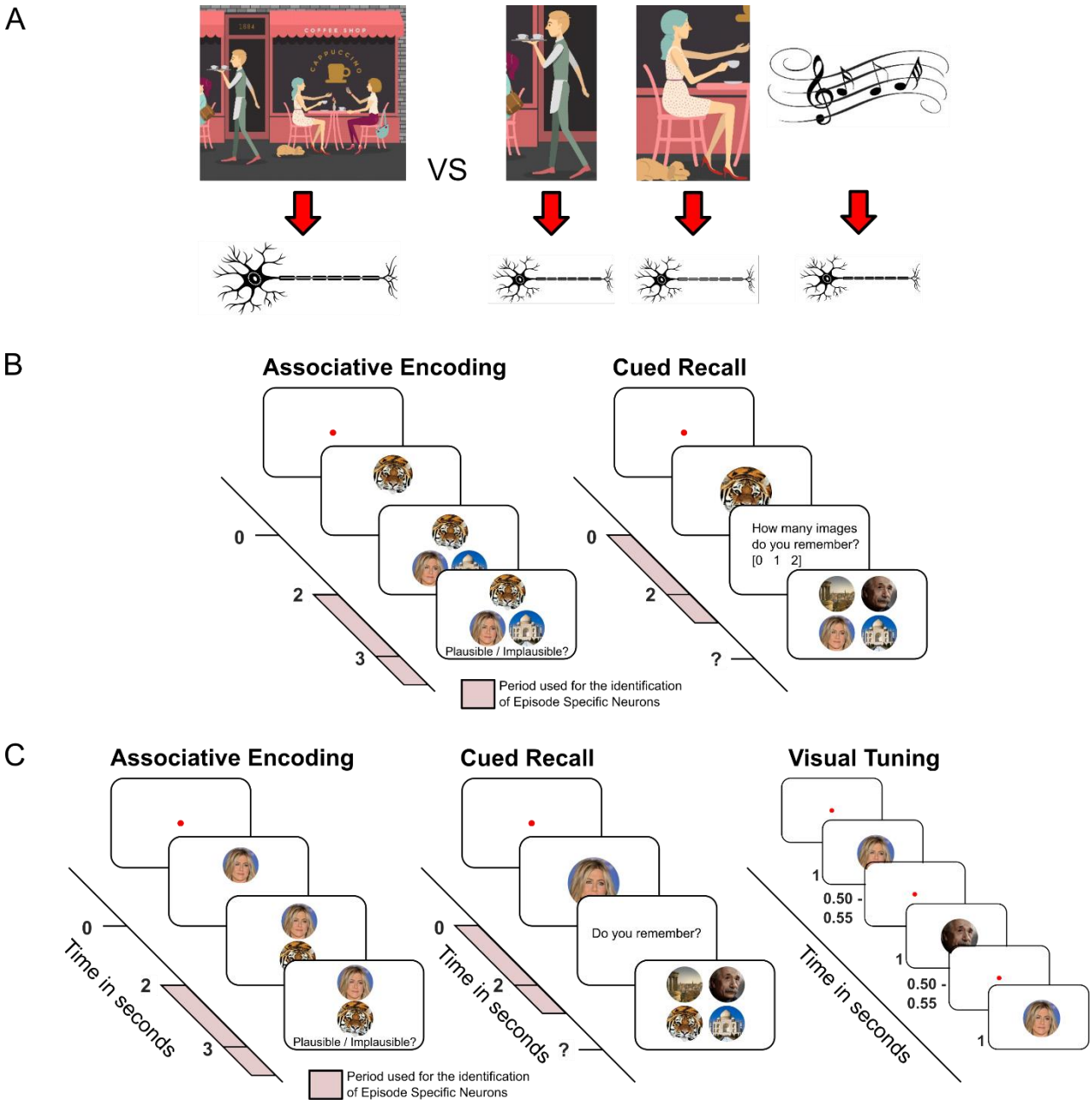


Figure 2.1. Difference between Indexing Theory and Concept Neuron based hippocampal coding of episodic memories, experiment procedure for experiments 1 & 2.

(A) Left: The classic Indexing Theory (Teyler & DiScenna, 1986) proposes that neurons in the hippocampus represent a conjunctive code that binds together all the elements that make

up the episode in the form of an index. Within this framework, neurons do not directly code for the elements themselves (i.e., the smell of the coffee, your friend, the background music, the café, etc.), but rather act as pointers to these different elements which themselves are coded elsewhere (i.e., the neocortex). Right: Some hippocampal neurons are thought to code for specific elements or concepts, which is why they are called Concept Neurons (Mormann et al., 2011; Mormann et al., 2008; Quian Quiroga et al., 2005). Within this framework, a group of neurons collectively code an episodic memory, with each neuron representing a specific element involved in that episode (i.e., a neuron coding for the coffee, another neuron coding for your friend, etc. (Quian Quiroga, 2020; Quiroga, 2012). It is important to note that one index or one concept is likely to be coded by an assembly of neurons, not a single neuron.

(B) Outline of the procedure for Experiment 1. During encoding, all participants were instructed to imagine a vivid episode involving an animal cue and two associate images (two faces, two places or a face and a place) and rated its plausibility. This approach is suitable for investigating episodic memory as originally defined by Tulving in 1972 (Tulving, 1972). During recall, participants were asked to retrieve the associated images when cued with the animal cue. The experiment was self-paced and every episode was learned and tested only once. Following each encoding block of roughly 20 episodes, participants performed a short distractor task. The pink areas represent the time windows used for subsequent analyses (see Methods).

(C) Outline of the procedure for Experiment 2. Left: The memory task was largely the same as in Experiment 1 (see Figure 2.1B). However, events consisted of one cue (either an animal, a face or a place) and one associate image (either an animal, a face or a place). Right: After the memory task, patients performed a visual tuning task where the previously used stimuli were shown multiple times in quick succession without a memory component. This approach has been traditionally used to identify putative Concept Neurons.

2.2 Materials and Methods

2.2.1 Procedure of memory Experiment 1

During the encoding phase of the experiment the participant associated a cue with two other stimuli. For each episode, the cue was a new picture of an animal. The stimuli could be pictures of either places, faces or both. Every picture was only shown once. Two seconds after the animal cue was presented, the associate stimuli were shown, while the animal cue remained on the screen. The participant was asked to create a vivid imaginary story involving

the cue and the two stimuli. This part of the experiment was self-paced. The task continued once the participant rated the plausibility of the imaginary story (plausible/implausible).
After the encoding phase, the participant performed a distractor task to rule out working memory effects. During the distractor task, participants had to indicate whether a random number (up to two digits) that appeared serially on the screen was odd or even. After each response, the participant received feedback indicating a correct or incorrect response. This task consisted of 15 trials.
During the retrieval phase, all cues from the previous encoding phase were presented sequentially in pseudorandom order. Each animal cue was presented for two seconds and subjects were tasked to retrieve the corresponding images. The participant was then asked how many associated images they remembered (none, one, or two). Participants had as much time to respond as they required. If the participant indicated that they remembered one or two images, they then were asked to select two pictures from an array of four pictures (two targets and two distractors that consisted of pictures from the previous encoding block which were associated with a different cue).
The experiment ended after the retrieval phase if the total runtime exceeded 40 minutes, or if the patient asked to abort the experiment. Otherwise, the experiment continued with the next encoding block. The encoding block initially consisted of 20 episodes but could be adjusted depending on the cognitive abilities of the patient. If the hit rate fell below 66.25%, fewer episodes were shown for the next block and vice versa if the hit rate surpassed 73.75%.
The patients performed the memory task on a laptop computer (Experiment 1: Toshiba Tecra W50, 60 Hz refresh rate; Experiment 2: Lenovo L390 Yoga, 60.01 Hz refresh rate), while either seated in a chair next to their bed or their hospital bed.

2.2.2 Procedure of memory Experiment 2

The second experiment is based on the first experiment with the following adaptations: participants are presented with one cue image (depicting an animal/place/face) and only one associate image (depicting an animal/place/face). During retrieval, participants were asked whether they remembered the associate image and the participants had to choose the correct associate from an array of four pictures (one target and three distractors that consisted of pictures from the previous encoding block which were associated with a different cue). The experiment was terminated upon request or when the runtime at the end of a retrieval block exceeded 30 minutes.

2.2.3 Visual tuning task procedure

For Experiment 2, the memory task was followed by a visual tuning task. During this tuning task, every image that was shown during the preceding memory task was displayed. Each image was shown six times in pseudorandom order on the screen for a duration of one second. The inter-image interval was jittered between 500ms and 550ms. To ensure attention, patients had to categorize the image as an animal, a place, or a face using the arrow keys on the keyboard.

2.2.4 Participants

For Experiment 1, eight patients were recorded in the Queen Elizabeth Hospital Birmingham (Birmingham, UK) (4 female; mean age: 36.25 years, from 26-49 years) and eight patients in the Universitätsklinikum Erlangen (Erlangen, Germany) (3 female; mean age: 36.125 years, from 26-53 years). For Experiment 2, 14 patients were recorded in the Universitätsklinikum Erlangen (Erlangen, Germany) (7 female; mean age: 33.857, from 19-58 years).

2.2.5 Ethical approval

Ethical approval was granted by the National Health Service Health Research Authority (15/WM/0219) and the Ethik-Kommission of the Friedrich-Alexander Universität Erlangen-Nürnberg (142_12 B). Informed consent was obtained in accordance with the Declaration of Helsinki.

2.2.6 Behavioural analysis

For the analysis of the first experiment, we considered an episode a hit if the participant correctly identified both stimuli. We considered an episode a miss if the participant either indicated not to remember any stimuli or did not remember either stimulus correctly. Participants correctly recalled on average 68.38% (*SE* = 4.64%) episodes in the first experiment (see Table S1) and on average 65.63% (*SE* = 4.45%) episodes in the second experiment (see Table S2). This is substantially more than would be expected by chance (16.7% and 25% respectively).

2.2.7 Statistical analysis

All statistical analyses were conducted using MATLAB R2020a on a computer running Windows 10 Enterprise. The significance threshold for all statistical tests was set at 0.05. Unless specified otherwise, all permutation tests were implemented with $N = 10{,}000$ random draws.

2.2.8 Co-Registering

For all but one patient, a pre-operational T1-weighted MRI scan was co-registered with a post-operational scan and normalized in MNI space using SPM12. For one patient, a post-operational CT scan was used instead of a post-operational MRI scan. Each microelectrode was localized either within the hippocampus, within the parahippocampus, or outside of both brain structures through visual inspection of an (see Figure A.1.). Only activity from microwires in Behnke-Fried electrodes assigned to the hippocampus was analysed in the main analysis of the current study. Neurons in the parahippocampus were analysed in an independent follow-up analysis.

2.2.9 Recording System and Electrodes

Patients were implanted with one to eight (see Table S1 and S2 for an overview) depth electrodes of the Behnke Fried type with microwire bundles (Ad-Tech Medical Instrument Corporation, USA) to localize epileptic foci. The electrode location was determined by clinical need. These single-use electrodes are made from platinum, have a diameter of 1.3mm and allow for simultaneous macro- and microcontact recordings. Platinum has a high impedance for lower frequency and a low impedance for higher frequency bands. As such it is suitable to pick up local extra-cellular action potentials. The micro contacts extended radially past the endpoint of the macro depth electrode, and each contained eight high-impedance microwires (~40-micron diameter) and one low-impedance microwire that is typically used for referencing.

The electrodes were connected to an ATLAS system (Neuralynx Inc, USA) consisting of CHET-10-A pre-amplifiers and a Digital Lynx NX amplifier and recorded with a sampling rate of either 32,000 Hz (Location: Birmingham) or 32,768 Hz (Location: Erlangen). Upon acquisition, an analogue bandpass filter from 0.1 Hz to 9,000 Hz was applied. In Experiment 1, microwires were referenced against another high impedance wire in two patients to increase the signal quality. All other microwires in Experiment 1 and Experiment 2 were referenced against the local low impedance wire.

2.2.10 Spike detection and spike sorting

In the following paragraph, we will outline the process used to filter the raw data, detect spike timestamps, extract features of the waveshape and cluster spike waveshapes into putative single neurons using the wave_clus toolbox. For a more in-depth description of the wave_clus algorithm, the reader is referred to (Chaure et al., 2018).

The unfiltered signal included both the local field potential and the action potentials of individual neurons. Action potentials are characterized by a very steep and transient amplitude in the signal. To extract these spikes, we first applied zero-phase filtering using a second-order bandpass elliptic filter in the range of 300-3,000 Hz. The resulting signal contained the information of the so-called spike band. Next, we segmented the continuous filtered data into epochs of five minutes. Segmenting the continuous data into smaller epochs had the advantage that noise in the signal did not increase the detection threshold for the whole recording and instead was limited to the segment in which it occurred (Chaure et al., 2018). Spike detection was performed separately for positive and negative deflections. Once a spike was identified, 64 data points around the spike maximum were extracted. This corresponds to a 2 ms window at a sampling rate of 32,000 Hz. The spike peak was aligned to the 20^{th} sampling point. To avoid misalignment of the spike, the waveshape was first up-sampled to 320 data points using cubic spline-interpolated waveforms and then down-sampled again (Chaure et al., 2018). Based on the extracted spike waveform, features were 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 were identified using a Lilliefors test and used for the clustering procedure (Chaure et al., 2018). Nonparametric clustering in the feature space was performed using superparamagnetic clustering (SPC). SPC grouped spike waves into clusters based on nearest-neighbour interactions (Blatt et al., 1996). Template-matching in Euclidian space was performed to assign unclassified waveforms to one of the identified clusters. The resulting clustering solution was then manually inspected and further optimized by rejecting artefact clusters, splitting clusters that represent multi-unit activity and merging clusters that likely stem from the same neural source. See Figure A.5. to Figure A.7. for an overview of the spike width, spike height, the Fano factor and the firing rate separately for ESNs and all other single units.

2.2.11 Identification of Episode Specific Neurons (ESNs)

For every single unit, we determined the number of spikes within each episode. During encoding, spikes from the onset of the associate images (two seconds after the cue onset i.e., when the whole information of the episode was present) until the end of the episode were considered. During the retrieval phase, spikes from cue onset until the time point at which participants indicated how many images they remembered were considered. We chose this time window because an episode could be reinstated following cue presentation, while after the response patients were presented with an array of images that could have potentially induced single-unit firing. Because the experiment was self-paced and longer episodes trivially contained more spikes, the firing rate (in hertz) was computed for each episode and single unit. In the next step, we z-scored this firing rate per single unit within all encoding episodes and retrieval episodes separately. Afterwards, we excluded all episodes that were later forgotten (for hit-ESNs) or that were later remembered (for miss-ESNs). Only sessions with at least eight episodes after this restriction were considered for further analysis. We then multiplied this standardized firing rate for encoding and retrieval episodes elementwise to gain an indicator for the reinstatement of firing for each episode (Figure 2.2). Alternative reinstatement measures are explored in the result section under *2.3.1. Identifying Episode Specific Neurons (ESNs)*, p. 28, and include (i) adding up the standardized firing rate between encoding and retrieval instead of multiplying them, (ii) increasing the minimum standardized firing rate from z = 1.645 to z = 2.6 and (iii) using a different reinstatement measure that normalizes the encoding and retrieval product by their absolute difference.

To estimate a threshold at which episode-specific firing reinstatement occurs on a single-unit level, we permuted the order of the encoding episodes and recomputed the elementwise product of the shuffled episode series. We repeated this permutation step 10,000 times and stored all output values. The 99^{th} percentile of these pooled values was then used as a threshold for firing reinstatement. As an additional constraint, z-scored firing during encoding and retrieval each had to exceed 1.645 ($\triangleq p_{\text{right-tailed}} < 0.05$) to make sure the elementwise product was not predominantly driven by a high firing rate in one of the two phases alone (i.e., either encoding or retrieval). This procedure is allowing us to threshold, but we do not have family-wise error corrected statistical significance at the single-unit level (there is no alpha inflation at the group level, see *2.2.12 Simulation of ESN identification*, p 24). Furthermore, we assume that single units fire independently. To ensure Concept Neurons tuned to the animal cue were not falsely interpreted as ESN activity, we excluded ESNs that showed a significant firing increase in response to the animal cue at encoding using the method described below under *Identification of putative Concept Neurons.*

In the second step, we calculated whether the number of ESNs (as identified in the above procedure) was above chance level. We did this by randomly choosing one of the permutations calculated in the first step for every single unit and checking whether it would be classified as an ESN under the same criteria outlined above. This approach is similar to a set-level effect in SPM (Penny et al., 2011). This process was repeated 10,000 times and the total number of single units which would be classified as an ESN in every single iteration of this process was used to build a distribution against which we compared our empirically discovered number of ESNs.

2.2.12 Simulation of ESN identification

We created a simulation using random pseudo-spike rates to determine whether our ESN analysis pipeline contains a bias towards false positive results. To create this simulation, we simulated the firing rate of 585 single neurons during 40 encoding and 40 retrieval trials by randomly drawing from a standard uniform distribution in the open interval of 0 to 1. These values were first multiplied by a variance factor that cycled from 2 to 5 and then z-scored independently for encoding and retrieval. Just as in the main ESN analysis we computed a reinstatement value for each trial by multiplying the two standardized synthetic firing rates. Next, we created a threshold by permuting the encoding and retrieval trial order 10.000 times while recomputing the shuffled reinstatement value. The 99th percentile was used as a threshold while the empirical standardized pseudo-firing rate had to be at least 1.645 during encoding and retrieval. If these criteria were met, we considered the neuron an ESN.

Then we computed the second-order (group level) permutation test by drawing a random first-order permutation for every single neuron and contrasted these values with the single neuron specific threshold. If the shuffled values satisfied the criteria for ESNs (i.e., encoding and retrieval standardized pseudo-firing rate at or above 1.645 and a reinstatement value above the neuron specific threshold) we considered the single neuron an ESN under the null distribution. By repeating this step 10,000 times we created a distribution under the H0 against which we could compare our initial random values. We repeated this entire process 1,000 times for each level of variance (2 to 5).

Because our initial pseudo-spikes were just random values, we expected 5% of all repetitions to yield a significant number of ESNs at any level of variance. If there was a bias, then more than 5% of all repetitions would contain a significant number of ESNs. As evidenced by Figure A.2. this was not the case for any levels of variance.

2.2.13 Identification of putative Concept Neurons

We have followed the method outlined in Mormann et al. (2011; 2008) to detect significant single-unit responses towards images. To this end, the 1000ms period after the stimulus onset was divided into 19 overlapping 100ms bins. The spike counts of each bin over all presentations of an image were compared to the 500ms baseline periods before stimulus onset for all images in the session using a two-tailed Mann-Whitney U test. We used the Simes' procedure to correct for multiple comparisons (Rødland, 2006). We performed this test twice, once with the commonly used threshold of $p < 0.0005$ and again with a liberal threshold of $p = 0.05$.

2.2.14 Identification of temporal Episode Specific Neurons (tESNs)

The analysis to identify neurons that showed a temporal firing reinstatement for specific episodes closely follows the outline described in *2.2.11 Identification of Episode Specific Neurons (ESNs)*, p. 23 – 24. For every neuron, we considered the spiking activity six seconds before until one second after the response during encoding and retrieval (the first and last second was later excluded to avoid edge artefacts).

We then convolved each spike with a gaussian kernel (standard deviation: 25ms/100ms/150ms, length: ± three standard deviations, peak normalized to one) creating a measure of instantaneous firing rate.

The main concern is that we do not know the ground truth of at what time point within a trial an episode was encoded or retrieved. To solve this problem, we cross-correlated the instantaneous firing rate during encoding with the instantaneous firing rate during the corresponding retrieval trial (maximum lag of +-2.5s). The maximum value of this sequence served as our empirical reinstatement value. We then shuffled the encoding and retrieval order and recomputed this reinstatement value 1,000 times. The 99th percentile of these values was used as a threshold. If the empirical reinstatement value reached this threshold, we considered the neuron a temporal Episode Specific Neuron (tESN). In the next step, for each neuron we randomly drew one of the permutations we calculated previously. Neurons whose permuted values reached or exceeded the threshold were considered tESNs under the null hypothesis. We repeated this process 1,000 times to build a null distribution against which we compared our empirical number of tESNs.

For Experiment 2, we further excluded all trials in which the given neuron showed a significant visual tuning using the methodology outlined under *2.2.13 Identification of putative Concept Neurons*, p. 25.

We tested the validity of this analysis by repeating the same analysis using random spike times. We generated these random spike times by first rounding the empirical spike times to the nearest integer and then drawing an equal number of pseudorandom integer values from a discrete uniform distribution between the first and last empirical spike times.

2.2.15 Firing rate spike convolution

To produce the visualisations in **Error! Reference source not found.**, we extracted spikes from one second before the cue onset until five seconds after cue onset for each episode. Binary spike times were convolved with a 251 ms Gaussian kernel (width factor: 2.5) to create a time-resolved signal of spike activity. We computed the average firing rate over time for all episodes (*ep*) during the baseline (*BL*) period 1,000 ms preceding the animal cue ($\overline{x}_{BL}$). We then z-scored the spike activity during the episode ($x_{ep,t}$) using the standard deviation ($std(\overline{x}_{BL})$) and mean ($\overline{\overline{x}}_{BL}$) across all pre-cue baseline periods (see equation (1)). To account for instances where no spiking activity occurred during the baseline period, 0.1 (see Ison et al., 2015) was added to the standard deviation ($std(\overline{x}_{BL})$). Episodes were then split into reinstated and non-reinstated episodes. Firing rates for each episode type (reinstated/non-reinstated) were then averaged over ESNs.

$$z_{ep,t} = \frac{x_{ep,t} - \overline{\overline{x}}_{BL}}{std(\overline{x}_{BL}) + 0.1} \quad (1)$$

2.2.16 Identification of Time Cells

We defined the beginning of an encoding block as the most salient event. Based on Umbach and colleagues (Umbach et al., 2020), we then extracted all spikes within each block and convolved them with a 251ms Gaussian kernel (width factor: 2.5). This created a block number x time points matrix. For our first analysis, we cut each encoding block into 40 equally sized bins, thereby normalizing block duration. We then used a Kruskal-Wallis test to determine whether any of the 40 bins significantly differed from each other.

We then performed a circular shifting permutation test to calculate whether we found a significant number of Time Cells. This is done by shifting a random number of values from the beginning of the vector to the end. This shifting was imposed on each block separately and repeated N = 10,000 times for every single unit. In a second test, the block length was determined by the longest block and shorter blocks were filled up with NaN values. This resulted in no normalization of time between blocks. The rest of the procedure is the same as described in the above paragraph.

2.3 Results

We analyzed recordings from two separate experiments (Experiment 1: 585 neurons in the hippocampus, 16 participants, 7 female; age mean = 36.125 years, from 26-53 years; Experiment 2: 216 neurons in the hippocampus, 14 participants, 7 female; age mean = 33.857 years, from 19-58 years) where patients were implanted with stereotactic Behnke-Fried depth electrodes in the hippocampus (Figure A.1), while they performed a memory association task (Figure 2.1B & Figure 2.1C).

During the encoding phase of Experiment 1 participants created a vivid mental story consisting of an animal cue and two associate images (two faces, two places or a face and a place). By contrast, Experiment 2 consisted of one cue and one associate image (both either an animal, face, or place). The encoding and recall phase of the experiment was interleaved with a short distractor task where patients had to judge whether a series of 15 numbers was odd or even. During the recall phase, the animal cue was presented again and participants were asked to retrieve the associate image(s). The experiments were self-paced and every episode was learned and retrieved only once. Participants correctly recalled on average 68.38% (*SE* = 4.64%) episodes in the first experiment (see Supplements Table S1) and on average 65.63% (*SE* = 4.45%) episodes in the second experiment (see Supplements Table S2). This is substantially more than would be expected by chance (16.7% and 25% respectively).

2.3.1 Identifying Episode Specific Neurons (ESNs)

For every neuron, we determined the firing rate during each episode at encoding and retrieval. We then z-scored the firing rate across all encoding and retrieval episodes and excluded all later forgotten episodes. This was done independently for encoding and retrieval to account for general differences in firing rates. We measured episode-specific firing reinstatement as the product of the standardized firing rates at encoding and retrieval (Figure 2.2A). Using a trial-shuffling procedure, we generated a distribution of reinstatement values expected by chance. A neuron was considered an ESN if (i) the empirical reinstatement value exceeded the 99th-percentile of the shuffled distribution for at least one episode and (ii) the standardized firing rate for encoding and retrieval of that episode each exceeded 1.645 ($\triangleq$ $p_{\text{right-tailed}} < 0.05$). The second criterion prevented the identification of ESNs which would excessively fire at only one phase of the task (i.e., encoding or retrieval).

It could be argued that ESNs identified in this manner could reflect the firing of cells tuned to the image of the animal cue, rather than the conjunction of all elements since the cue is

episode-unique and presented during encoding and retrieval. To address this issue, in Experiment 1, we excluded neurons that showed a significant firing increase during the first second after the encoding of the animal cue for episodes that were later reinstated (see Methods). This procedure has traditionally been used to identify putative concept neurons (Mormann et al., 2008; Mormann et al., 2011; Quiroga et al., 2005). Using this approach, we identified a significant number of hippocampal ESNs in Experiment 1 (136 out of 585 neurons ≙ 23.25%; $p < 0.001$; permutation test; Figure 2.2B). Comparable results are obtained when (i) adding up the standardized firing rate between encoding and retrieval instead of multiplying them (Enc & Ret >= 1.645, reinstatement: Enc + Ret) (125 ESNs; $p < 0.001$), (ii) increasing the minimum standardized firing rate from z = 1.645 to z = 2.6 (Enc & Ret >= 2.6, reinstatement: Enc * Ret) (29 ESNs; $p < 0.001$) and (iii) using a different reinstatement measure that normalizes the encoding and retrieval product by their absolute difference (Enc & Ret >= 1.645, reinstatement: (Enc * Ret) / |Enc - Ret|) (53 ESNs; $p < 0.001$). This reinstatement measure has the important advantage of considering the similarity between the encoding and retrieval firing rates.

In Experiment 1 117 out of 136 ESNs (≙ 86.03%) coded for a single episode. Two example ESNs are shown in Figure 2.3. These ESNs are unlikely to be Concept Neurons tuned to the animal cue as the firing rate during encoding reaches its maximum only after the presentation of the associate stimulus (see **Error! Reference source not found.**A).

It is of note that the proportion of neurons that can be classified as ESNs is proportional to the number of events learned and retrieved (the same is the case for Concept Neurons). This is because we apply the threshold derived from the first permutation test to all episodes, without family-wise error correction. As such it is not suitable to determine the sparseness of the hippocampal code. However, the proportion of ESNs of all recorded neurons is useful as an estimation of how many ESNs we can expect in future analyses.

It is crucial to understand that this alpha-level inflation does not extend to the group-level permutation test, where the same number of tests are applied to randomly shuffled data. We have added a simulation using random values as spike rates to show that there is no inflation of the alpha error at the group-level at which we interpret our findings (see Methods; Figure A.2).

ESNs are suggested to reflect a unique coding mechanism of the hippocampus (Teyler and DiScenna, 1986; Teyler and Rudy, 2007). In line with this, we did not find a significant number of ESNs in the parahippocampus (15 out of 104 neurons, $p = 0.5396$; permutation test). However, it should be noted that all 104 parahippocampal neurons originate from only five different microwire bundles over 14 sessions in five different patients, and therefore

these results, albeit in line with the indexing theory, should be interpreted with caution. The second experiment did not contain a significant number of ESNs in the parahippocampus (3 out of 25 neurons, $p = 0.1199$).

To conclude, we find a significant number of ESNs in the hippocampus, but not in the parahippocampus. The analysis approach we use to identify ESNs is robust to deviations in the parameter space.

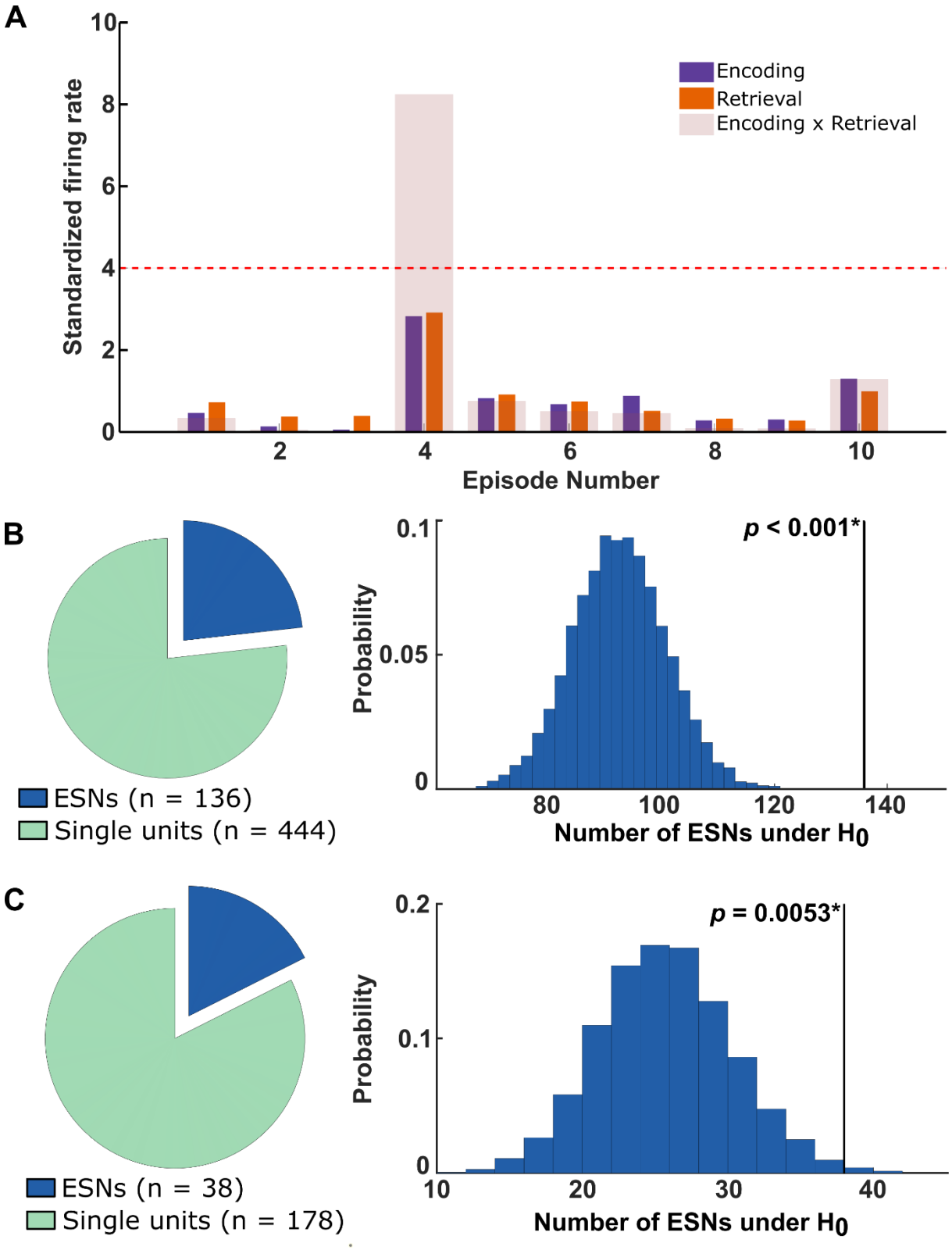


Figure 2.2. Analysis schematic and number of ESNs identified in experiments 1 & 2.
(A) A schematic for identifying Episode Specific Neurons (ESNs) is shown. The diagram shows the z-scored firing rate on the y-axis for ten simulated episodes on the x-axis colour-coded for encoding and retrieval (purple and orange, respectively). The transparent bars encompassing encoding and retrieval indicate the product of encoding and retrieval firing

rates, which is used as the measure of episode-specific firing reinstatement. The dotted red line shows the threshold (derived from a shuffling procedure, see Methods). Because of the way ESNs are defined, they are required to fire substantially above their average firing rate during encoding and retrieval, which rules out neurons that generally show an increased firing rate during remembered episodes.
(B) Identified ESNs during Experiment 1. Left: The pie chart depicts the number of ESNs that show significant firing reinstatement to at least one episode (dark blue) and the number of neurons that showed no firing reinstatement (green). Right: The number of ESNs as expected by chance and the empirical number of ESNs (136 out of 585 neurons; $p < 0.001$; permutation test).
(C) Same as (B) but for Experiment 2. Out of a total of 105 hippocampal neurons, we identified 38 ESNs ($p = 0.0053$; permutation test).

2.3.2 ESNs do not code for the content/visual properties of the cue or associate image

Traditionally, visually responsive neurons have been identified using the repeated presentation of a stimulus. In the above analysis, we only present the animal cue once, which is suboptimal for ruling out Concept Neurons tuned to the animal cue. To ameliorate this shortcoming, in Experiment 2 we added a visual tuning task (Figure 2.1C) after the memory association task. During the visual tuning task, images from the memory task were repeatedly shown in quick succession. This approach is widely used to identify putative Concept Neurons that respond to one of the images independently of any memory processes (Mormann et al., 2011; Mormann et al., 2008; Quian Quiroga et al., 2005). When excluding Concept Neuron activity in this independent dataset, we replicated our previous results and identified a significant number of ESNs (38 out of 216 neurons ≙ 17.59%; $p = 0.0053$; permutation test; Figure 2.2C). In Experiment 2 34 out of 38 ESNs (≙ 89.47%%) coded for a single episode.
However, traditional Concept Neuron detection methods might be too conservative to identify weakly tuned Concept Neurons. To address this concern, we drastically reduced the threshold of what constitutes a Concept Neuron, i.e., lowering the uncorrected threshold from $p = 0.0005$ to $p = 0.05$, which increased the number of Concept Neurons from 58 to 155 (out of 216 neurons). Remarkably, incorporating this liberal threshold to exclude potential Concept Neurons, had little effect on the number of ESNs which remained almost unchanged (36 out of 216 neurons ≙ 16.67%; $p = 0.0025$; permutation test).

During a typical tuning task an average of 108.7 (min: 80; max: 156) different images are shown and each image is tested for visual tuning. There is no correction for multiple testing rendering a threshold at $p < 0.05$ very liberal.

It is conceivable that some images that are presented during the visual tuning task act as cues that reactivate some ESNs. These reactivated ESNs would then be erroneously rejected as Concept Neurons. However, in practice, only four potential ESNs were excluded based on the visual tuning task (six when lowering the Concept Neuron threshold). We suspect that ESNs were not reactivated during the visual tuning task because the images were shown only for one second. This might be sufficient to cause Concept Neuron firing, but too short to elicit episodic memory retrieval. Nonetheless, we cannot rule out that in some cases ESNs were reactivated during the visual tuning task and subsequently rejected. However, this would only make our analysis for identifying ESNs more conservative.

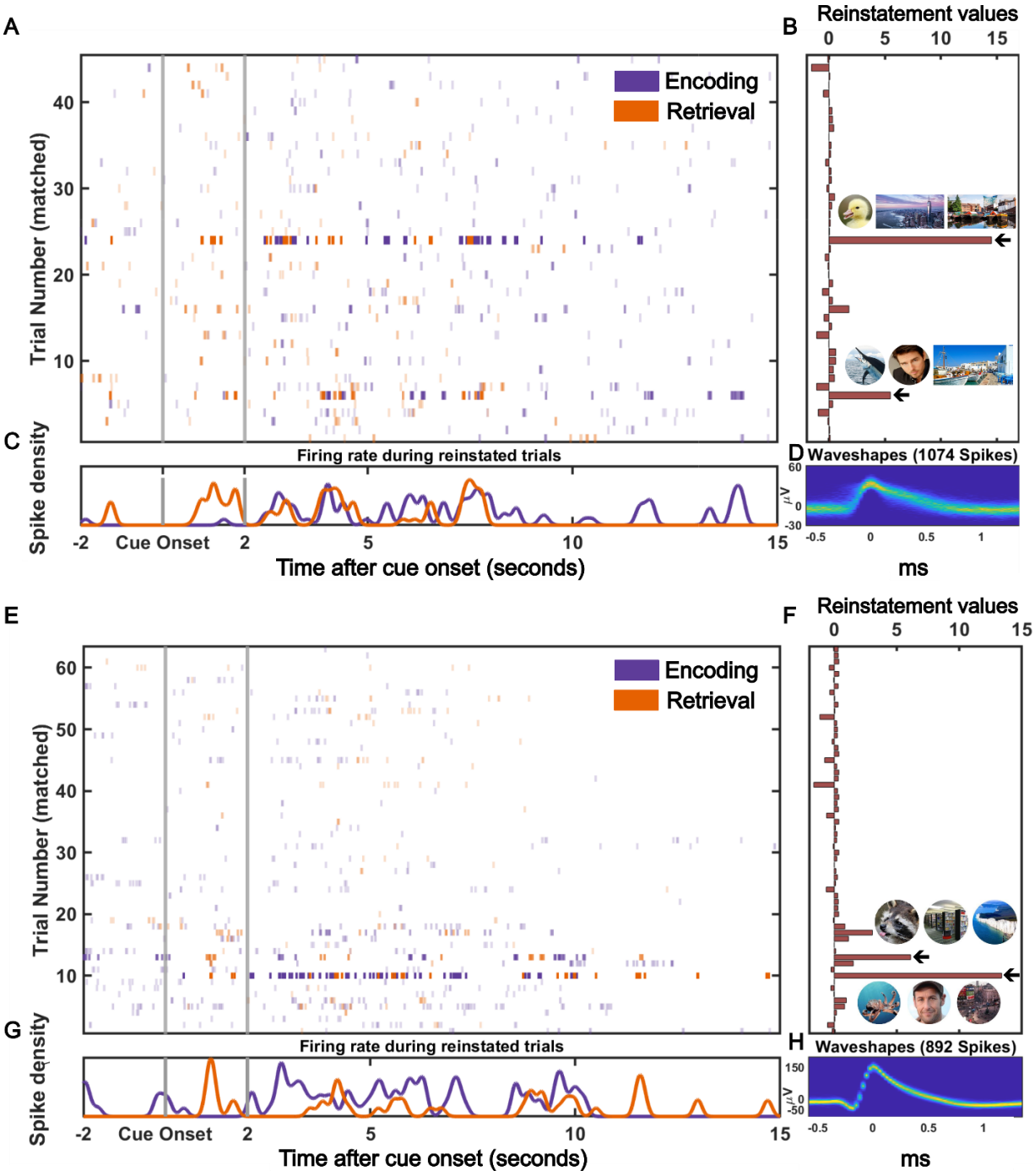


Figure 2.3. Firing patterns for two example Episode Specific Neurons.

(A) Spike raster plot. Each line indicates a spike. On the x-axis is time and on the y-axis are the episodes. Color-coded in purple for encoding and orange for retrieval. The transparency is adjusted according to the reinstatement values in that specific episode.

(B) Reinstatement values and the animal cues with the respective associate images for reinstated episodes (indicated by the black arrows).

(C) Spike density plot for reinstated episodes. Note that the experiment is self-paced and episode length varies.

(D) 2D histogram of the waveshape of that particular unit (Niediek et al., 2016).

(E-H) same as (A-D) but for a different example ESN.

2.3.3 ESNs are limited to later remembered episodes

We have so far demonstrated that ESNs reinstate their firing rate when remembering a unique episode. This reinstatement cannot be explained by the semantic content or visual properties of the used image, which strengthens the notion that ESNs code for memories. In line with this, we did not find a significant number of ESNs when limiting our analysis to later forgotten episodes (15 out of 585 neurons ≙ 2.56%; $p = 0.4229$; permutation test). However, this result could stem from a lower number of forgotten events (see Table S1). To counter this bias, we equalized event numbers between later remembered and later forgotten events for every neuron by randomly sampling (with replacement) later remembered events as many times as participants forgot an event. If any of the sampled events were later reinstated, we considered this neuron a miss-ESNs under the null hypothesis. By repeating this procedure 10,000 times we generated a distribution of how many miss-ESNs were expected if the number of later remembered and later forgotten events were equal. This analysis did not result in a significantly lower empirical number of miss-ESNs compared to hit-ESNs ($p = 0.7032$, bootstrapping test). To conclude, we did not find a significant number of ESNs when restricting our analysis to episodes that were forgotten. However, when considering that fewer episodes were forgotten than remembered there was no difference in the number of hit-ESNs and miss-ESNs.

2.3.4 Identification of temporal Episode Specific Neurons (tESNs)

The previous identification of ESNs relied on a rate code, i.e., the standardized mean firing rate during one episode at encoding and retrieval. We have adapted this analysis to identify neurons that reinstate a temporal pattern of firing. For every neuron, we considered the spiking activity six seconds before until one second after the response during encoding and retrieval (the first and last second was later excluded to avoid edge artefacts).

By convolving each spike separately with different gaussian kernels (standard deviations: 25ms/100ms/150ms, length: three standard deviations, peak normalized to 1) we created a measure of instantaneous firing rate. Because we do not know the exact times when an episode is encoded or retrieved, we cross-correlated this trial-specific instantaneous firing rate during encoding and retrieval and considered the maximum value as the reinstatement value. We repeated this process after shuffling the encoding and retrieval trial order 1,000 times and took the 99th percentile as a threshold for the empirical reinstatement value. If the empirical reinstatement value reached this threshold, we considered the neuron a temporal Episode Specific Neuron (tESN; Figure A.4). In the next step, we randomly drew for each

neuron one of the previously calculated permutations. If these permuted values reached or exceeded the threshold the neuron was considered a tESNs under the null hypothesis. We repeated this process 1,000 times to build a null distribution against which we compared our empirical number of tESNs. We found a significant number of empirical tESNs in Experiment 1 when using a smoothing kernel of 25ms (292 out of 585 neurons; $p < 0.001$), 100ms (280 out of 585 neurons; $p < 0.001$), and 150ms (280 out of 585 neurons; $p < 0.001$). For Experiment 2, we further excluded all trials in which the given neuron showed a significant visual tuning (see Methods). With this additional constraint, we found a significant number of tESN in Experiment 2 when using a smoothing kernel of 25ms (100 out of 216 neurons; $p = 0.003$), 100ms (103 out of 216 neurons; $p < 0.001$), and 150ms (103 out of 216 neurons; $p < 0.001$).

We then tested the validity of this analysis using random spike times. We generated these random spike times by first rounding the empirical spike times to the nearest integer and then drawing an equal number of pseudorandom integer values from a discrete uniform distribution between the first and last empirical spikes times. Independent of the gaussian kernel and experiment we did not find a significant number of tESN (all $p > 0.2$).

In conclusion, we show in two separate experiments a significant number of neurons that reinstate an event-specific temporal firing pattern during successful memory retrieval.

2.3.5 ESNs do not code for time

Recent studies in humans show that some hippocampal neurons code specific time points invariant across repetitions, which are referred to as Time Cells (Reddy, Zoefel et al., 2021; Umbach et al., 2020). We investigated whether our dataset contains such Time Cells (TC) using a similar method as employed by (Umbach et al., 2020). Due to the self-paced nature of our experiment, each encoding block varied in length. To accommodate this, we used both the unaltered block length, as well as a normalized block length within one recording session (see Methods). Of all 585 recorded cells, 12 (normalized) and 9 (non-normalized) fulfilled the criteria of TCs, which is below chance level (p values > 0.9; permutation test). Critically, there was no significant overlap between neurons that behaved like TCs and ESNs (p values > 0.3; permutation test) indicating that ESNs cannot be construed as TCs.

2.3.6 ESNs show a wider waveshape than other neurons

We found some evidence that spike waveshapes of ESNs are wider than those of other units (Figure A.5A; $p = 0.0563$; with data from Experiment 1 and $p = 0.0121$ with data from both

experiments; both independent samples t-test), possibly indicating that ESNs are physiologically different from other neurons. In the hippocampus, a wider waveshape has previously been associated with excitatory cells (Prestigio et al., 2019), therefore suggesting the ESNs are predominantly excitatory neurons. There was no significant difference in the spike height or Fano factor between ESNs and other neurons (unpaired t-tests; all *p* values > 0.3; Figure A.5B and A.6).

2.3.7 Recorded neurons are mostly single neurons and not multi-units

Although we tried to separate multi-units into single neurons as best as possible during the spike sorting procedure (see Methods), some units might still represent activity from multiple neurons. We thus employed the method outlined by Tankus and colleagues (Tankus et al., 2009) to classify units into single units and multi-units, using the inter-spike interval and spike waveshape variability as objective criteria. In the first experiment, 373 out of 585 units (≙ 63.76%) were classified as single units (95/136 ESNs ≙ 69.85%), while in the second Experiment 132 out of 216 units (≙ 61.11%) were classified as single units (20/38 ESNs ≙ 52.63%). If we limit our analysis to neurons that satisfy these stringent criteria for putative single neurons, we still find a significant number of ESNs in the first experiment 95 out of 373 single neurons; $p < 0.001$), but not for the second experiment (21 out of 132 single neurons; $p = 0.071$).

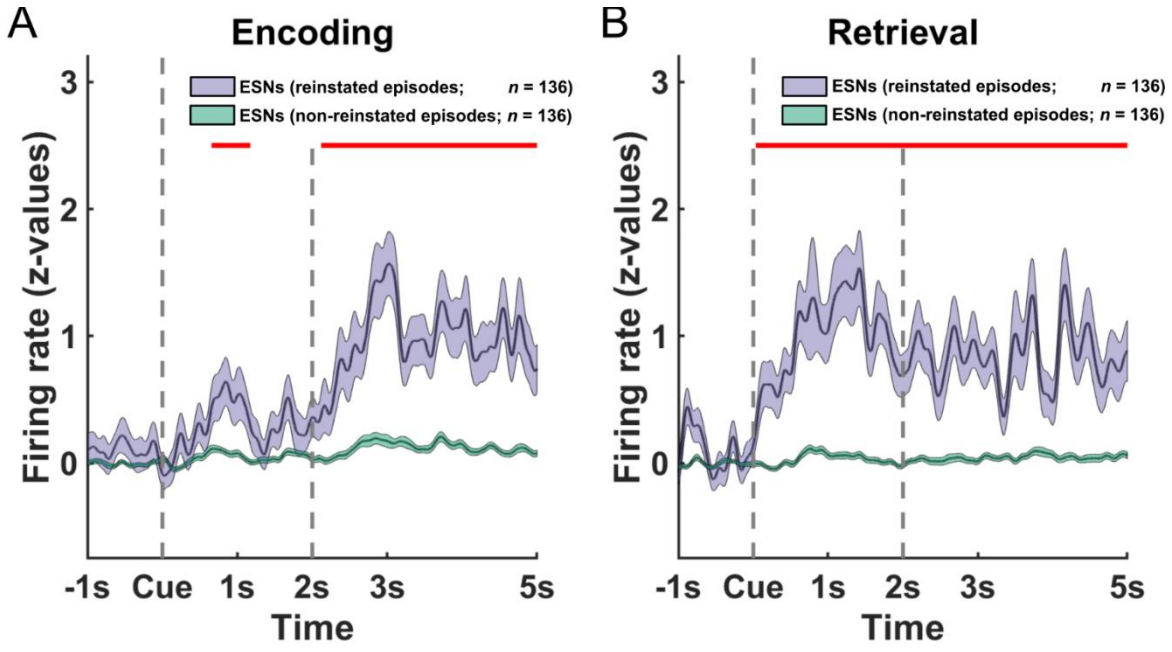


Figure 2.4 Firing rate of ESNs during reinstated (purple) and non-reinstated (green) episodes.
(A) Firing rate of ESNs from cue onset until five seconds later during memory encoding. The red line marks time points where the average ESN firing rate during reinstated episodes

(n = 136) significantly exceeds the firing rate during average non-reinstated episodes (n = 136; cluster permutation test; Maris & Oostenveld, 2007).
(B) Same as (A) but for memory retrieval. The shaded areas indicate the SEM.

2.4 Discussion

Using an associative episodic memory paradigm in human epilepsy patients, we identified hippocampal neurons that are active during the initial encoding of a unique episode and later reinstate their firing rate when successfully remembering the same episode. Therefore, we term these neurons Episode Specific Neurons (ESNs). The activity of these neurons could not be explained by a firing rate increase towards specific images or time points. We verified these results using a number of alternative reinstatement measures and changes in the hyperparameter space.

Previous studies have demonstrated that Concept Neurons increase their firing rate during memory retrieval when the image they are tuned to is part of the memory (Gelbard-Sagiv et al., 2008; Ison et al., 2015). We used two approaches to ensure that the ESNs we identified are not Concept Neurons that selectively respond to visual elements or semantic concepts: (i) in Experiment 1 we excluded ESNs that were visually responsive to the presentation of the animal cue at encoding. (ii) Following the episodic memory task in Experiment 2 patients completed a visual tuning task using all previously presented stimuli. This is a standard method to identify putative Concept Neurons (Ison et al., 2015; Mormann et al., 2011; Mormann et al., 2008; Quian Quiroga et al., 2005) and allowed us to exclude episodes where a neuron showed a visual tuning to either the cue or the associate image. Using this approach, we replicated our results from Experiment 1 in a new sample of patients and found a significant number of ESNs while also verifying that these neurons do not selectively respond to visual elements or semantic concepts. Importantly, this finding was robust even when dramatically reducing the threshold of what constitutes a Concept Neuron. Taken together these analyses reinforce the argument that ESNs are memory-related.

The existence of ESNs does not exclude Concept Neurons from playing a role in episodic memory processes. Concept Neurons might code the semantic aspect of an episode (i.e., the general concept of "coffee shop"). However, according to the Indexing Theory (Teyler and DiScenna, 1986; Teyler and Rudy, 2007), hippocampal neurons that perform this indexing function should have no initial tuning and are allocated to a specific episode during memory formation (i.e., the coffee shop in a specific setting). The behaviour of ESNs would be

consistent with such an indexing function and may add crucial event-specific information to an episode, that Concept Neurons cannot encode themselves.

We found a significant number of ESNs when excluding potential multiunits in the first experiment. However, we could not replicate these findings in the second experiment. This was likely because this restriction resulted in too few single neurons in the second experiment.

Because we do not know the exact time points when episodes are encoded or retrieved, we used a rate code approach in the first instance for these analyses (i.e., averaging the number of spikes over a time of interest and encoding and retrieval). In addition, we present preliminary evidence for a reinstatement of a temporal firing code which we uncovered by shifting the instantaneous firing rate (i.e., the spike times convolved with a gaussian kernel) using a cross-correlation.

The Indexing Theory proposes that this coding mechanism is unique to the hippocampus. In line with this, we did not find a significant number of ESNs in the parahippocampus. However, these findings are based on a relatively small sample size and should be considered preliminary. Future studies are needed to ascertain the regional specificity of ESNs to the hippocampus.

We did not find a significant number of ESNs when restricting our analysis to later forgotten episodes. However, there was no significant difference between the number of ESNs when considering later remembered and later forgotten events. Hippocampal neural reinstatement might occur without behavioural memory retrieval. This could be due to downstream processing being disrupted (i.e., due to interference or selective attention). Another possible explanation for this finding is that in some cases during memory encoding patients created an episodic memory that did not incorporate the presented associate stimuli. While retrieval would lead to neural reinstatement, the patients would not be able to choose the correct associate images.

Time Cells (TC) are neurons that invariantly fire at specific, reoccurring time points (Reddy, Zoefel et al., 2021; Umbach et al., 2020). We did not find a significant number of TCs in our study and there was no significant overlap between TCs and ESNs. This might be because the self-paced nature of the task introduced too much time variation between too few learning blocks to uncover TC dynamics. However, the absence of TCs in our paradigm corroborates ESNs as independent from TCs.

We found tentative support that ESNs have a wider waveshape than other neurons. This suggests that ESNs are likely excitatory cells (Prestigio et al., 2019). Alternatively, it is possible that ESNs and neurons with a narrower waveshape are located in different

hippocampal subfields. Unfortunately, with the current methods, we lack the precision to designate neurons to individual subfields (Quiroga, 2019).

One limitation of the current study is that every event was encoded and retrieved only once. However, the very nature of episodic memories is one-shot learning and the ability to subsequently perform mental time travel. Any neural substrate that supports this function must occur after a single bout of learning and subsequent retrieval of a single episode. Our method honours this fundamental characteristic which is the defining feature of episodic memory as originally stated by Endel Tulving (Tulving, 1972). Arguably, a repeated design would have allowed for a more reliable ESN identification. However, each memory reactivation leads to a transient plasticity of the memory trace until it is reconsolidated again. During this time window profound changes in the neurons that code for the initial memory trace might occur (Nader & Hardt, 2009). To avoid this potential confound, every episode is learned and retrieved only once in the present experiments. The stability of ESNs over repeated reactivations and extended periods, therefore, remains an interesting topic of research for future studies.

Our results are consistent with previous studies using fMRI that have shown item-specific activity reinstatement in the hippocampus (Chadwick et al., 2010; Mack & Preston, 2016) where similar representations are associated with distinct activity patterns (Bakker et al., 2008; Berron et al., 2016). These findings are suggestive of an episode-specific neural code, which is consistent with our results. However, due to the coarse resolution of fMRI, these previous results cannot disambiguate whether this event-specific code is driven by a population of event-specific concept neurons, or whether it is driven by a population of event-specific indexing neurons. We here provide evidence for the latter.

Previous intracranial work has identified a multitude of different neurons that detect novelty or familiarity (Rutishauser et al., 2006; Rutishauser et al., 2008; Rutishauser et al., 2015) as well as episode boundaries and event onsets (Zheng et al., 2022). These cell types generally fired to many episodes, whereas the vast majority of identified ESNs in our experiments coded a single episode. When quantifying neural firing reinstatement between scene encoding and recognition, recent work relied on population activity (i.e., considering the activity of all recorded neurons) (Zheng et al., 2022). In contrast, we showed here that neural reinstatement takes place on the level of a single neuron. Importantly, we expect that an episode is coded by an assembly of ESNs from which we sampled only one due to the limited number of neurons that can be recorded with the currently available methods. These findings are in line with previous work showing that episodic memories in the hippocampus are coded in a sparse distributed way (Wixted et al., 2018; Wixted et al., 2014). However, there are

various reasons why we refrain from making any claims regarding sparsity in the present study. A neural code can be sparse in two ways (Wixted et al., 2018). A neural code can be population sparse, which is the case when a low percentage of neurons respond to a given stimulus. It can also be lifetime sparse, which refers to a low percentage of stimuli that a given neuron responds to.

On the one hand, we artificially induce lifetime sparsity (and by extension population sparsity) because we (i) standardize the firing rate during encoding and retrieval and then (ii) multiply these two values. On the other hand, we drastically reduce the sparsity because we test for reinstatement at each episode without correcting for multiple comparisons. It is very important to understand that while this leads to alpha-level inflation at the level of the neuron, this does not extend to the group-level at which we interpret our findings. We have confirmed that our analysis does not have a bias towards positive findings using a simulation (see Figure A.2). Unfortunately, that also means that in the present study we have to refrain from making any claims regarding lifetime sparsity. Moreover, in the current study, the associated image was not shown on the screen during memory retrieval. This mental reinstatement is a core feature of episodic memory which is difficult to assess with recognition-based memory paradigms.

In conclusion, we found neurons in the hippocampus that show firing reinstatement in response to a specific conjunction of elements within a unique episode. These Episode Specific Neurons did not fire in response to individual concepts (Concept Neurons) or to specific, re-occurring time points (Time Cells). We propose that during memory formation an assembly of ESNs acts as a pointer or index that initially binds the elements of an episode together, in line with the Indexing Theory (Bowman & Wyble, 2007; Teyler & DiScenna, 1986; Teyler & Rudy, 2007). Reactivation of this pointer allows ESNs to reinstate the episodic memory previously encoded. Importantly, because ESNs reinstate unique episodes, they contain a time and content component. However, rather than reflecting the underlying coding mechanism, this time and content aspect necessarily emerges from the conjunctive code of an episode that is unique in content and time.

Chapter 3

High frequency power reinstatement in the human hippocampus for specific episodic memories

Abstract

Previous work has identified single neurons in the human hippocampus that significantly increase their firing rate during the encoding and retrieval of specific episodic memories (Episode Specific Neurons; ESNs). High frequency power (40-200 Hz; HFP) in the local field potential has been used as a proxy for multi-unit activity. We here studied the reinstatement of HFP in the hippocampus of patients while they completed a memory association task. Consistent with earlier observations we find a significant number of microwires that show a reinstatement of HFP from encoding to retrieval in individual episodes. Importantly, this reinstatement is not driven by a content-specific code (i.e., population activity of Concept Neurons). This effect is limited to later remembered episodes and not present for later forgotten episodes. These findings extend the discoveries of the previous chapter from the single neuron level to the population activity reflected in the local field potential.

3.1 Introduction

Episodic memories refer to the memory of distinctive events composed of multiple, multimodal elements that occurred at a specific time and space (Tulving, 1972, Tulving, 2002). In the previous chapter, we investigated the formation and retrieval of these episodic memories at the level of single neurons in the human hippocampus. These neurons (called Episode Specific Neurons; ESNs) increase their firing rate during encoding and retrieval of specific episodic memories. We provided compelling evidence that this episode specific code is separate from Concept Neurons. In this chapter, we will delve into the neurophysiological substrates of memory processing that is one level above individual neurons: the local field potential (LFP). In contrast to local neural firing, LFPs reflect the aggregate of a myriad of local and distant transmembrane currents (Buzsáki et al., 2012). We will focus on the role of high frequency power (HFP; 40-200 Hz) as a proxy of local synchronous spiking activity (Buzsáki et al., 2012; Manning et al., 2009; Nir et al., 2007; Ray et al., 2008). Most of the literature examining the relation of spiking activity and HFP is based on studies in monkeys in early sensory cortical areas that have a topographic structure (Buzsáki et al., 2012; Leszczyński et al., 2020; Ray et al., 2008; Ray & Maunsell, 2011; Whittingstall & Logothetis, 2009), but some evidence has been reported in humans (Kucewicz et al., 2014; Manning et al., 2009; Miller et al., 2009; Nir et al., 2007). Although neighbouring neurons in the hippocampus are not structured topographically and often represent very different concepts (De Falco et al., 2016; Redish et al., 2001) there is some evidence that the HFP-spiking relationship remains intact (Manning et al., 2009).

It is unclear if enough neurons are part of one assembly of ESNs to increase HFP, and further if these neurons are close enough in space and fire in synchrony. Preliminary evidence comes from Rutishauser and colleagues who reported that roughly 10-20% of all neurons in the hippocampus and amygdala responded to novel stimuli (Rutishauser et al., 2006; Rutishauser et al., 2008), which is likely enough to elicit HFA. However, the authors do not report whether these neurons respond to specific new episodes or new episodes in general and how many of them reinstate their firing rate during retrieval. Based on the average number of identified Concept Neurons, recorded neurons, and presented images, it is estimated that approximately one million neurons within the medial temporal lobe code for a given concept. This represents only 0.1% of the total number of neurons in the MTL (Quian Quiroga, 2012), which likely does not impact HFP.

In conclusion, we postulate a reinstatement of power in the high frequency band from encoding of specific trials to their reinstatement during an episodic memory task. As Concept

Neurons are thought to be part of smaller assemblies (Quian Quiroga, 2012) we expected not to find changes in high frequency power induced by specific concepts.

3.2 Materials and Methods

3.2.1 Common methods

For a description of the *experimental procedures*, the *participants*, ethical approval, *behavioural analysis*, *co-registering of the MRIs, recording system and electrodes*, *spike detection and spike sorting*, and *Identification of Episode Specific Neurons (ESNs)* please see the methods section of Chapter 2 (p. 18 – 24).

3.2.1 Statistical analysis

All statistical analyses were conducted using MATLAB R2020a on a computer running Windows 10 Enterprise. The significance threshold for all statistical tests was set at 0.05. Unless specified otherwise, all permutation tests were implemented with N = 1,000 random draws.

3.2.2 LFP pre-processing

We downsampled the LFP data from microwires that contained neurons in the hippocampus to 1,000 Hz and applied a fourth-order Butterworth bandstop filter with a centre frequency of 50 Hz (± 1 Hz) and its harmonics up to 300 Hz, to remove line noise.

3.2.3 LFP Artefact Rejection

For each microwire, we computed the bandpass-filtered signal between 40 Hz and 200 Hz using a first-order Butterworth filter. We identified any data points exceeding five standard deviations from the mean of this signal as artefacts and excluded the one-second intervals preceding and following them.

3.2.4 Identification of Episode Specific Microwires (ESWs)

We considered neural activity from the onset of the associated image to the patient's response in encoding trials, and from the cue onset to the response onset in retrieval trials. To account for edge artefacts, we extended these trial definitions by 100ms on each side. We then performed a wavelet analysis using wavelets from 40 Hz to 200 Hz in steps of 5 Hz and a

width of 7 cycles, on the linenoise-removed broadband signal. After removing all artefacts (see *3.2.3 LFP Artefact Rejection*, p. 45), we computed the mean power over all frequencies. Trials that consisted of 50% or more artefacts during encoding or retrieval were excluded, and if fewer than nine trials remained, the microwire was not considered for further analysis. We z-scored the remaining HFA power values independently for encoding and retrieval, and afterwards excluded later forgotten trials. Finally, we defined the element-wise product of the encoding and retrieval standardized HFA power as a proxy for episode-specific reinstatement. To calculate a threshold for this episode-specific firing reinstatement we permuted the order of the encoding and retrieval episodes and recomputed the reinstatement value. We repeated this step 1,000 times and took the 99th percentile as a threshold against which we compared the empirical reinstatement value. If the empirical reinstatement exceeded the threshold and its standardized power at encoding and retrieval was at least 1.645 ($\triangleq p_{\text{right-tailed}} < 0.05$), we considered this microwire an Episode Specific Microwire (ESW). This procedure allows for thresholding but does not correct for multiple comparisons on the level of a microwire. To determine whether there was a significant number of microwires that showed an episode-specific power reinstatement, we randomly drew one of the previously calculated permutations for each microwire and determined whether it would be classified as an ESW under the same criteria as before. In each of the 1,000 permutations, we summed up the number of shuffled ESW which we then used to create a null distribution against which we compared the empirically determined number of ESW. To generate Figure 3.3, we repeated the time-frequency analysis in the range of 3 Hz and 200 Hz in 50 logarithmically spaced steps for all microwires that exhibited a HFP reinstatement in at least one episode. For each ESW we calculated the mean HFP during reinstated and non-reinstated episodes and then averaged the respective power spectra across all ESW. To determine the statistical significance of the results, we used a cluster-based permutation test (Maris & Oostenveld, 2007).

3.2.5 Identification of putative Concept Specific Microwires (CSWs)

We have adapted the method created by Mormann and colleagues (Mormann et al., 2008; 2011) for detecting Concept Neurons to identify microwires whose HFP was reliably increased following the presentation of a specific image. For each microwire, we divided the local field potential of the 1000ms interval post-stimulus into 19 100ms overlapping bins, with the 500ms preceding stimulus onset as the baseline period. To prevent edge artefacts, we extended the testing and baseline intervals by 100ms on either side. We performed a time-frequency analysis using wavelets in the range of 40 Hz to 200 Hz (stepsize: 5 Hz) and

a width of 7 cycles, allowing us to estimate the time-resolved power. We then averaged the power over all frequencies and within each time bin. If more than one of any of the six repetitions of an image contained over 50% artefacts that time bin was discarded for all repetitions. We then compared the mean HFA power in the remaining 19 bins across all six presentations of an image with the mean HFA power of all baseline periods in the session using a Mann-Whitney U test. We corrected for multiple comparisons using the Simes' procedure (Rødland, 2006). To test whether our dataset has a significant number of CSWs for each microwire we shuffled the trial order and recomputed the CSW detection pipeline. We repeated this step 1,000 times to generate a distribution of how many CSW to expect under the null hypothesis.

3.2.6 Correlation between HFP and spiking activity

After pre-processing the LFP of the microwire on which a neuron was recorded (see *3.2.2 LFP pre-processing*, p. 45) we segmented the data into later remembered episodes. During memory encoding the time of interest started at the onset of the associate image(s) and ended when the patient gave their response. In contrast, during memory retrieval, the time of interest started at the cue onset and ended when the patient gave their response. We added 100ms on each side to account for edge artefacts. Then, we performed a wavelet analysis between 40 Hz and 200 Hz in steps of 5 Hz and a width of 7 cycles, and averaged the power across frequencies. We then normalized the HFP across time, using a z-transformation, and concatenated this standardized power values across all episodes. To compute the instantaneous firing rate of the corresponding neuron, we convolved the firing times with a Gaussian kernel (kernel parameters: mu = 0, standard deviation = 50ms, length = 300ms, normalized peak to 1). We then z-scored this instantaneous firing rate and concatenated all episodes. Subsequently, we performed a linear correlation between the concatenated standardized HFP and the concatenated standardized instantaneous firing activity, separately for encoding and retrieval. To assess the statistical significance of the correlation we shuffled the data circularly and recomputed the correlation with this shuffled data. We repeated this step N = 10,000 times and compared the empirical correlation coefficient with the resulting null distribution of shuffled correlation coefficients. We performed this analysis twice: once for neural activity during reinstated episodes and once for all other episodes.

3.3 Results

We conducted two different experiments in which patients implanted with stereotactic Behnke-Fried depth electrodes completed a memory association task (see Figure 2.1) In Experiment 1 we recorded from 1011 microwires and 585 neurons in the hippocampus (16 participants, 7 female; average age = 36.13 years, from 26-53 years) and in Experiment 2 we recorded from 344 microwires and 216 neurons in the hippocampus (14 participants, 7 female; average age = 33.86 years, from 19-58 years). During the encoding phase of Experiment 1 patients were instructed to mentally form a vivid story containing an animal cue and two associated images (two faces, two places, or one of each). Experiment 2 only had one associate image and either cue or associate could be a face, a place or an animal. After the encoding phase a short distractor task commenced during which patients had to determine whether a series of 15 numbers were odd or even. During the retrieval phase, the cue image was presented and the patient was asked to retrieve the associated image(s). Each episode was learned and retrieved once and the experiment was completed at the participants' own speed.

3.3.1 Reinstatement of high frequency power

To investigate high frequency power reinstatement, we calculated the average power within a range of 40 Hz to 200 Hz in steps of 5 Hz for every microwire. During encoding we considered neural activity from the time point the associated image was presented until the patient gave their response. During retrieval the time of interest stretched from the cue onset to the response. We z-scored the power values independently for encoding and retrieval and subsequently excluded episodes that were later forgotten. We defined the element-wise product of the standardized encoding and retrieval power values as a measure of episode-specific reinstatement. Using a trial-shuffle procedure we re-computed these reinstatement values 1,000 times. If any empirical reinstatement value exceeded the 99th percentile of these permuted values and if the standardized power at encoding and retrieval during that episode exceeded a value of at least 1.645 we considered this microwire an Episode Specific Microwire (ESW; see Figure 3.1 for an example). To estimate how many ESW we can expect by chance we then randomly drew one of the previously calculated permutations for each microwire and applied the same thresholding technique to these shuffled reinstatement values. This allowed us to create a distribution of ESW under the null hypothesis against which we could compare the number of empirically identified ESW. Using this approach, we found a significant number of ESW in Experiment 1 (n = 144 out of 1010 microwires, p

= 0.0310; permutation test; see Figure 3.2). However, there was no significant number of ESW when limiting the analyses to later forgotten episodes (p = 0.305; permutation test). We subsequently contrasted the power spectra of reinstated episodes with non-reinstated episodes from 3 Hz to 200 Hz using 50 log-spaced frequency points. A cluster-based permutation test revealed that during reinstated trials, the power was significantly increased from 9.9 Hz to 200 Hz (p < 0.001) at encoding and from 15.3 Hz to 200 Hz (p < 0.001) at retrieval (see Figure 3.3).

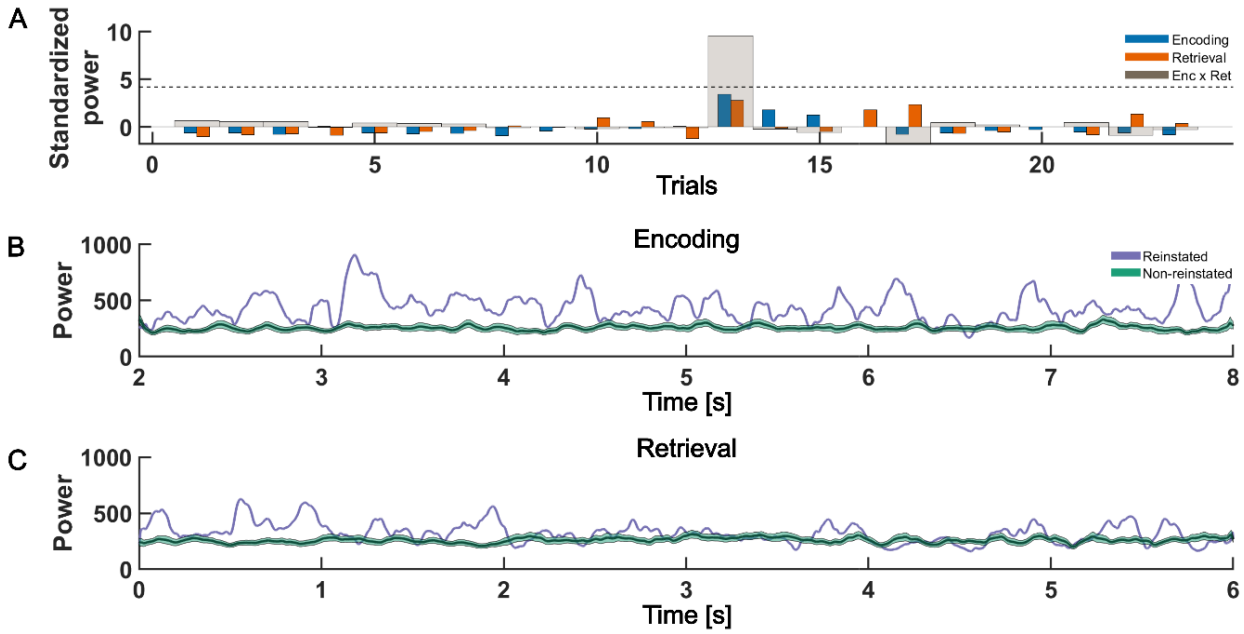


Figure 3.1. Example Episode Specific Microwire (ESW).

(A) The bar plots show the z-scored HFP on the y-axis for 23 episodes on the x-axis colour coded for encoding (blue) and retrieval (orange). The transparent bar encompassing the standardized HFP represent their element wise product, which is used as a measurement for episodic memory reinstatement. The dotted line represents the threshold which is calculated based on a permutation test.

(B) The time resolved HFP (y-axis) during memory encoding with the time in seconds (x-axis) starting from the associate image onset for reinstated episodes (purple) and non-reinstated episodes (green). The shaded area represents the SEM.

(C) Same as (B), but during retrieval and starting at the cue onset.

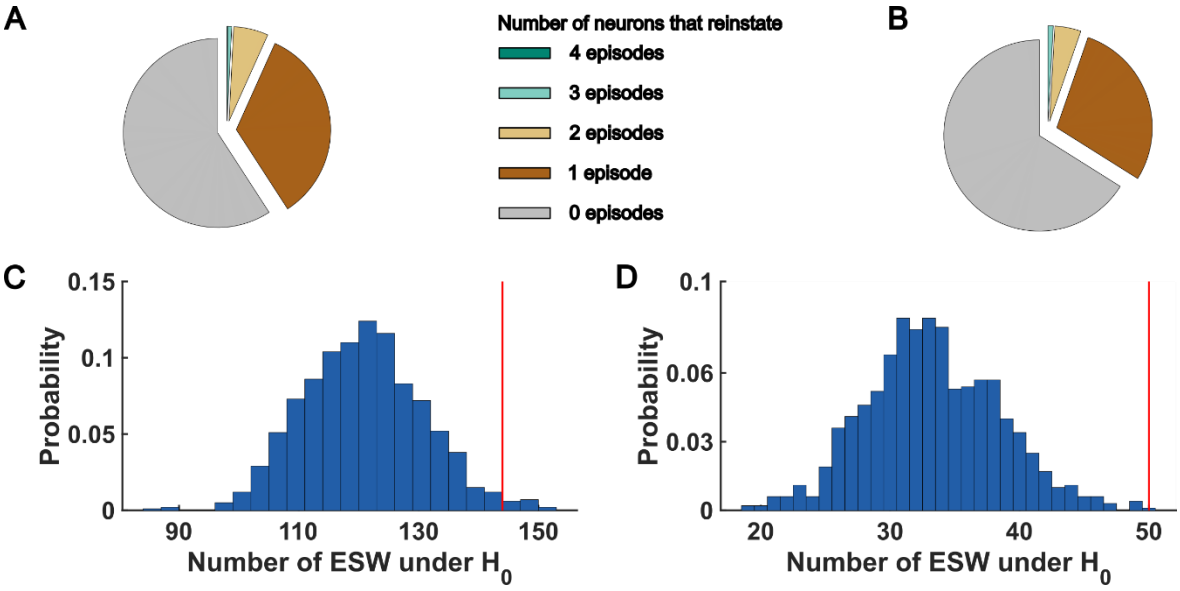


Figure 3.2. Number of reinstated episodes and number of ESW expected under the null hypothesis.

(A) Pie chart showing the number of episodes each neuron reinstated during Experiment 1 (zero episodes: 598 microwires; one episode: 345 ESWs; two episodes: 60 ESWs; three episodes: 7 ESWs; four episodes: 1 ESWs).

(B) Same as (A), but for Experiment 2 (zero episodes: 224 microwires; one episode: 97 ESWs; two episodes: 15 ESWs; three episodes: 3 ESWs).

(C) Distribution of the number of ESWs expected by chance and the number of empirically found ESW (red line) in Experiment 1.

(D) Same as (C) but for Experiment 2.

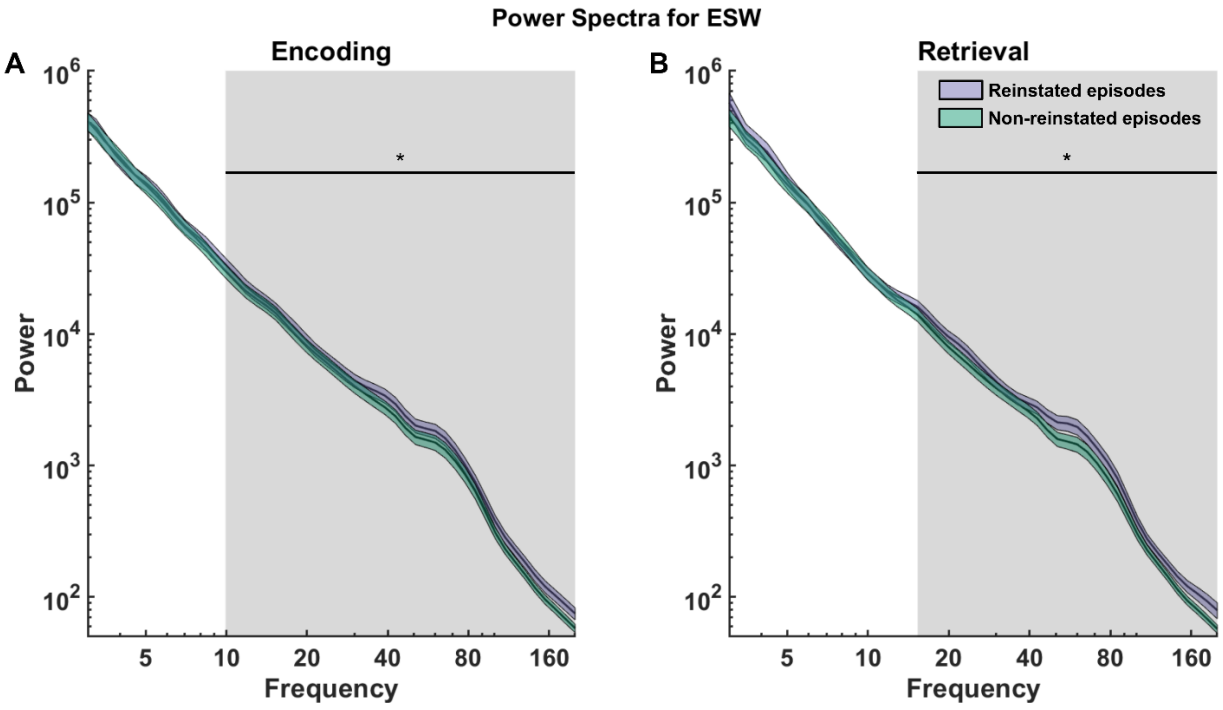


Figure 3.3. Power spectra for reinstated episodes (purple) and non-reinstated episodes (green) during (A) encoding and (B) retrieval.

The x-axis displays the frequency, ranging from 3 Hz to 200 Hz in 50 logarithmically spaced increments. The y-axis displays the power on a logarithmic scale to enhance editorvisibility. The shaded regions show the SEM. The grey rectangles specify frequencies at which the power during reinstated episodes significantly exceed the power of non-reinstated episodes (Maris & Oostenveld, 2007).

3.3.2 HFP reinstatement is not content dependent

The second experiment included a visual tuning task, during which the same images that were used in the preceding memory task were presented repeatedly without an episodic memory component. This approach has been traditionally used to detect neurons responding to specific concepts or categories (Mormann et al., 2008; Quian Quiroga et al., 2005) and allowed us to exclude all episodes that contained an image which reliably evoked a HFP increase during a visual tuning task. We defined Concept Specific Microwires (CSW) as any microwire with a significant increase of HFP in any of 19 overlapping 100ms time bins following the image presentation across all six repetitions in comparison to a 500ms pre-stimulus baseline period using a Mann-Whitney U test (see Methods). We carried out the analysis twice, once with the typically used cut-off threshold of $p = 0.0005$ and again with a more liberal cut-off threshold of $p = 0.05$. Note that no corrections were made for testing multiple images for tunings, thus making a threshold of $p = 0.05$ very liberal. No CSWs were detected at $p = 0.0005$; however, when the threshold was lowered to $p = 0.05$, we found a

significant number of CSWs (86 out of 344 microwires, p = 0.005, permutation test). Because no CSWs were detected at a cut-off of p = 0.0005, no episodes were excluded in the ESW analysis. In Experiment 2 we replicated our prior results and found a significant number of ESWs (n = 52 out of 344 microwires, p = 0.003). We then repeated the ESW analysis, this time excluding episodes with significant CSW activity at a threshold of p = 0.05. Despite this threshold change, we identified a significant number of ESWs (n = 50 out of 344 microwires, p = 0.001). Of note, although the more liberal CSW threshold led to the identification of fewer ESWs the resultant p-value is lower. This is because the reduced threshold is also applied when determining the number of permuted ESWs (i.e., ESWs expected by under the null hypothesis).

In summary, we discovered a memory code in the form of a HFP reinstatement between encoding and retrieval of individual episodes across two independent experiments. Although we were unable to detect any CSW activity using the traditionally used threshold, we detected a significant number of CSWs with a more liberal threshold. Importantly, our findings could not be accounted for by a content-specific code (i.e., CSWs).

3.3.3 HFP correlates with ESN and single neuron firing

Next, we examined the correlation between HFP and single neuron firing in our sample. We first determined the instantaneous firing rate of each ESN during reinstated episodes. In a separate analysis we calculated the instantaneous firing rate during non-reinstated episodes. We segmented the LFP data into later remembered episodes and performed a wavelet analysis from 40-200 Hz. For each episode we averaged the power in that frequency range. We then z-scored the instantaneous firing rate and the HFP estimate across time. Finally, we concatenated each episode separately for encoding and retrieval. We performed a linear correlation between the standardized HFP and the standardized instantaneous firing rate and assessed the statistical significance by comparing it with the correlation values that we obtained through circular shuffling. In Experiment 1, HFP and ESN firing during reinstated episodes correlated with r = 0.132 during encoding (r^2 = 0.017, p < 0.001; permutation test) and r = 0.120 during retrieval (r^2 = 0.014, p < 0.001; permutation test). Firing during non-reinstated episodes significantly correlated with HFP during encoding (r = 0.110, r^2 = 0.012 p < 0.001; permutation test) and during retrieval (r = 0.083, r^2 = 0.007, p < 0.001; permutation test). A similar relationship was found in Experiment 2, where HFP and firing during reinstated episodes correlated with r = 0.140 at encoding (r^2 = 0.020, p < 0.001; permutation test) and r = 0.135 during retrieval (r^2 = 0.018, p < 0.001; permutation test). Firing during non-reinstated episodes correlated with HFP with r = 0.072 during encoding

(r^2 = 0.005, p < 0.001; permutation test) and r = 0.063 during retrieval (r^2 = 0.004, p < 0.001; permutation test).

3.4 Discussion

Episodic memories refer to distinctive events that occurred at a specific time and space. These memories are composed of multiple components. In Chapter 1 we identified how the human hippocampus processes these episodic memories. These neurons (called Episode Specific Neurons; ESNs) increase their firing rate during encoding and retrieval of specific episodic memories. In the present chapter, we extended these findings from single neurons to the population level by investigating the local field potential (LFP) as a proxy of multi-unit activity. We analysed two independent datasets that were collected using microelectrodes located in the human hippocampus while patients performed a memory association task. Power in the high frequency band (40-200 Hz) on a significant number of microwires was reinstated from encoding to retrieval of specific episodes. These findings cannot be explained by a content code (i.e., HFP induced by the presence of particular concepts). Applying the traditional criterion used to detect Concept Neurons seems to be too conservative to detect significant increases in the HFP. However, when lowering this threshold, we found a significant number of microwires that show a consistent HFP increase when presenting specific concepts (CSW) despite the relatively small assembly size of Concept Neurons (Quian Quiroga, 2012). Importantly, the same threshold was also lowered for the group-level permutation test, which we used to determine the number of CSW expected under the null hypothesis. Concept Neuron activity might be reflected in the HFP due to the spatial clustering of Concept Neurons within the vicinity of a microwire (but see De Falco et al., 2016; Redish et al., 2001). Alternatively, multiple Concept Neurons coding the same concept might be active within a short delay, leading to a higher deflection in the LFP. The development of new electrodes (such as Durand et al., 2022; Dutta et al., 2019; Jun et al., 2017; Paulk et al., 2022) that enable the recording of multiple neurons that code the same concept at more precisely known locations might offer an answer to this question.
Our analyses revealed that the power differences between reinstated and non-reinstated episodes exceeded the frequency range of 40-200 Hz that we used to differentiate the two. Reinstated episodes were characterized by an increased power from 10 Hz (during encoding) and 15 Hz (during retrieval), implying that the distinction between reinstated and non-reinstated episodes may not be limited to 40-200 Hz, but could be attributed to either an offset or a spectral tilt of the 1/f power spectrum. Future studies will need to carefully

disentangle the individual contributions of oscillatory changes, a power offset, and a spectral tilt between reinstated and non-reinstated trials.

The range of high frequency activity (40-200 Hz) overlaps with the so-called ripple band (80-120 Hz). Excitatory input from CA3 induces ripple activity in CA1, which is characterized by highly synchronized neural firing (Buzsáki, 2015). Ripple activity has been linked to memory consolidation and replay of previous experiences (Jadhav et al., 2012; Roux et al., 2017; Vaz et al., 2019). For example, research by Vaz and colleagues (Vaz et al., 2019), has implicated coupled ripple activity to coordinate the information flow between the MTL and the temporal cortex (also see Ngo et al., 2020). Considering this, future studies should investigate the extent to which the here reported HFP memory reinstatement effect is driven by activity in the ripple range. Here we reported a significant correlation between neural firing and the HFP on the microwire on which the neurons were recorded. The correlation coefficient was low, explaining roughly 0.4-2% of the variance. It is important to note, that we correlated the firing rate of individual neurons with the HFP, which includes the activity from disproportionally more neurons with heterogeneous firing rates. We, therefore, expect a higher correlation when analysing more neurons. Moreover, it is conceivable that different microwires from the same bundle may be better suited to pick up transmembrane currents from nearby neurons. Additionally, there is substantial variance in the literature on which frequency range constitutes the high frequency band (Ray et al., 2008: 60-200 Hz; Whittingstall & Logothetis, 2009: 30-100 Hz; Leszczynski et al., 2020: 70-150 Hz; Nir et al., 2007: 40-130 Hz). Future studies should aim to identify which frequencies are most indicative of single neuron firing taking into consideration differences between neuron types (excitatory or inhibitory cells). Furthermore, a computational model suggested that synchronous firing is more influential in increasing HFP compared to firing alone (Ray et al., 2008). Unfortunately, due to the limited number of single neurons that can be recorded using currently available microwires we cannot resolve this question. Recording more neurons using newer electrodes may enable us to disentangle the roles of synchrony and firing in relation to HFP in the future. A larger brain coverage would also allow the investigation of what role the recorded brain area plays in moderating the relationship between neural firing and HFP (Leszczyński et al., 2020).

To conclude the present chapter, we identified a significant number of microwires which show a HFP reinstatement during encoding and retrieval of specific memories (ESW) consistently across two independent datasets. This HFP activity showed a low, but significant correlation with neural firing of ESNs and single neurons during encoding and retrieval. Although we did not find reliable HFP increases to specific concepts using the

traditionally used threshold, we identified a significant number of concept coding microwires using a more liberal threshold (CSW). Importantly, the HFP reinstatement for specific memories could not be attributed to this content code. Taken together the present work extends findings from the level of the single neuron and provides a potential link to surface EEG recordings (Buzsáki et al., 2012).

Chapter 4

Absence of periodic theta increase and no theta spike-field coupling in the human hippocampus during episodic memory processing

Abstract

Theta oscillations play a central role in memory processing. Recent findings point towards there being not one dominant theta frequency in the human hippocampus, but rather two: a slow theta (2-5 Hz) and a fast theta (5-9 Hz) oscillation. It has been suggested that successful memory processing is reflected in a narrowband theta increase as well as a 'tilt' in the aperiodic power spectrum, where lower frequencies are diminished and higher frequencies increased. Furthermore, according to an influential theory memory encoding and retrieval occurs in opposite theta phases so newly encoded memories do not cause catastrophic interference with older memories.

We investigated these hypotheses in two independent samples of intracranial microwire recordings. Contrary to previously reported findings, our results provide inconclusive evidence regarding narrowband slow and fast theta power and an aperiodic tilt. Our research did not reveal consistent evidence that neurons that increase their firing rate during encoding and retrieval of specific episodes (Episode Specific Neurons; ESNs) or other non-tuned neurons fire at a distinct theta phase during encoding and retrieval. Likewise, we found no significant theta phase difference between neurons firing at encoding and retrieval.

4.1 Introduction

In the preceding chapters, we found evidence of an episode specific neurophysiological marker at both the single-neuron level and in the high frequency power of the microwire local field potential (LFP). Next, we will explore another prominent frequency range in the hippocampus, the theta oscillation, and investigate how it relates to single-neuron spiking.

Research in the role of theta oscillations on learning on memory go back to the late 70s (Berry & Thompson, 1978; Winson, 1978). Winson (Winson, 1978) showed that lesioning the medium septum caused a reduced hippocampal theta rhythm along with an impaired spatial memory. In line with this, higher theta power in rabbits was associated with augmented learning (Berry & Thompson, 1978). Since then, evidence regarding the role of theta oscillations in episodic memories has been contradictory. While most studies employing surface EEG report increases in theta power, most iEEG studies report a memory induced theta power decrease (Herweg et al., 2020). Herweg and colleagues (Herweg et al., 2020) suggested that this might be because studies frequently contrast later remembered with later forgotten memories and therefore conflate domain-general cognitive processes, such as attention and perception, with memory-specific processes. Because domain-general cognitive processes are assumed to lead to a spectral tilt (i.e., less low frequency power and more high frequency power), a narrow band theta power increase induced by memory processing might be obscured. To ameliorate this shortcoming researchers should not contrast successful memory with unsuccessful memory but instead should compare strength of memory (e.g., retrieval confidence, amount of detail in contextual retrieval, retrieved spatial distance to encoded location in a navigational task; Herweg et al., 2020) . Another reason how surface EEG might show a theta power increase, although the LFP shows a decrease is if theta over larger areas synchronizes but decreases in amplitude. The decrease is truthfully reflected in the LFP, but activity on the scalp is integrated over larger areas and thus more synchronous theta could lead to higher scalp theta power (Herweg et al., 2020). Taken together these considerations imply theta activity as an integral part of memory processing and suggest that conflicting evidence arises due to different recording methods (EEG/iEEG), memory contrasts (success vs success or vs failure) and frequency ranges (broadband vs narrowband).

An increased narrowband theta activity is in line with the prediction from a computational model and theoretical considerations that theta synchronization in the hippocampus is necessary for memory processing (Hanslmayr et al., 2016; Parish et al., 2018). More recent findings in humans demonstrated that behavioural response times in memory tasks are

modulated by theta oscillations (ter Wal et al., 2021) and that theta binds together the multiple elements within an episode (Clouter et al., 2017; Griffiths et al., 2021; Roux et al., 2022).

A central requirement of the hippocampus is the ability to encode new information without interfering with related previous experiences. Hasselmo and colleagues developed a computational model that solves this conundrum by moving encoding and retrieval processes to opposing phases in the theta rhythm (Hasselmo et al., 2002). Empirical support for this 180° shift between memory encoding and retrieval has been recently found by Kerrén and collegues (Kerrén et al., 2018; Kerrén et al., 2022). The relation between single neuron firing and ongoing theta oscillation contains more information than the neural firing alone (Huxter et al., 2003; Jacobs et al., 2007). Place cells in the hippocampus are neurons that code for specific spatial locations. As rodents move towards a location, a place cell fires at increasingly earlier phases of the ongoing theta oscillation. One can therefore decode the position of the rodent in relation to a place by combining the theta phase and the neural firing (O'Keefe & Recce, 1993). In humans, a stronger spike-theta coupling (Rutishauser et al., 2010) as well as neurons locking to faster theta oscillations (Roux et al., 2022) predicts successful memory. Importantly, recent findings suggest that there are two distinct theta rhythms governing the human hippocampus: a slow (2-5 Hz) and a fast (5-9 Hz) oscillation (Goyal et al., 2020; Kota et al., 2020; Lega et al., 2012).

We therefore hypothesized that (i) later remembered episodes show a shift in the aperiodic power spectrum and an accompanying increase in oscillatory fast and slow theta power in comparison to later forgotten episodes. (ii) We also expected this change in aperiodic and periodic activity to manifest when comparing episodes in which ESNs reinstate their firing rate (as described in Chapter 2) and episodes which are not reinstated. (iii) We hypothesized that neurons, particularly ESNs, fire at distinct slow and fast theta phases during the encoding and retrieval of episodic memories, and that there is a substantial phase offset between encoding and retrieval.

4.2 Materials and Methods

4.2.1 Shared Methods

For a description of the *experimental procedures*, the *participants*, ethical approval, *behavioural analysis*, *co-registering of the MRIs*, *recording system and electrodes*, *spike detection and spike sorting*, and *Identification of Episode Specific Neurons (ESNs)* please see the methods section of Chapter 2 (p. 18 – 24).

4.2.2 Statistical analysis

All statistical analyses were conducted using MATLAB R2020a on a computer running Windows 10 Enterprise. The significance threshold for all statistical tests was set at 0.05. Unless specified otherwise, all permutation tests were implemented with N = 1,000 random draws.

4.2.3 Periodic and aperiodic theta analysis

To investigate periodic (i.e., oscillatory) and aperiodic activity (i.e., 1/f activity), we first downsampled the LFP in every microwire to 1000 Hz and bandpassed the signal using a fourth order Butterworth filter at 50 Hz ±1 Hz and harmonics up to 300 Hz. An episode was labelled as reinstated if any neuron on the respective microwire contained a single neuron that showed a significant firing increase during encoding and retrieval (i.e., an ESN; see Chapter 1). We defined the time of interest as the period two seconds prior to the response at memory encoding and retrieval. In Experiment 1, an episode was considered correctly remembered if the patient correctly chose two out of two associate images and labelled as forgotten if the patient indicated they do not remember any associates or if they chose no correct associate.

For each episode, we extracted the periodic and aperiodic part of the signal using the FOOOF implementation (Donoghue et al., 2020) in Fieldtrip (Oostenveld et al., 2011) in a frequency range from 1 Hz to 200 Hz. We analysed two contrasts of the periodic and aperiodic activity: (i) reinstated episodes against non-reinstated episodes in microwires with ESNs, and (ii) correctly remembered episodes against forgotten episodes (excluding reinstated episodes).

For the periodic analysis, we averaged activity within the slow (2-5 Hz) and fast (5-9 Hz) theta bands and then conducted paired-sample t-tests to compare oscillatory activity between contrasts and one-sample t-tests to test for significant oscillatory activity. For the aperiodic analysis we performed paired-sample t-tests between contrasts, with the offset and tilt as dependent variables.

4.2.4 Theta components and pre-processing

As a first step, we downsampled the microwire signal to 100 Hz. Because we do not know the relative position of the recorded neurons to the microwires within a bundle of electrodes by extension we do not know if the microwire on which the neuron was recorded best represents the neural input into the neuron. For this reason, we took into consideration all

eight microwires and generated two theta components using generalized eigendecomposition (Cohen, 2017). The generalization of the eigendecomposition extends the eigendecomposition to a case with two square matrices. For an eigenvalue decomposition with a singular square matrix, the eigenvector with the highest eigenvalue accounts for the maximal variance in the underlying square matrix and is pairwise orthogonal to the other eigenvectors. In contrast, the eigenvector with the highest eigenvalue in a generalized eigendecomposition can be understood as the filter that maximizes the difference between the two input matrices. The eigenvectors in a GED are independent, but not orthogonal. In practice when applied to two covariance matrices where one matrix represents the broadband activity and the other matrix is generated using a narrowband signal the first eigenvector yields a spatial weighting that maximizes the narrowband activity and minimizes the broadband activity. This eigenvector can be applied to the narrowband filtered multichannel data to generate a narrowband component (Cohen, 2017; see Figure 4.1 for a five second data segment).

Based on previous literature (Goyal et al., 2020; Kota et al., 2020; Lega et al., 2012) we computed a slower theta component in the frequency range of 2 Hz to 5 Hz and a second, faster component in the range of 5 Hz and 9 Hz. To generate these components, we first applied a first order Butterworth filter to bandpass the broadband signal in all eight microwire channels between 2 Hz and 5 Hz (slow theta component) or 5 Hz and 9 Hz (fast theta component). We then demeaned the signal and computed a covariance matrix using this narrowband signal, which we divided by the number of samples. Next, we computed a second covariance matrix using the entire broadband signal. We computed the generalized eigendecomposition of these two covariance matrices and used the eigenvector with the highest eigenvalue as a spatial filter for the narrowband filtered signal to generate a narrowband component. We then applied the Hilbert transform to the narrowband component to get the analytic signal.

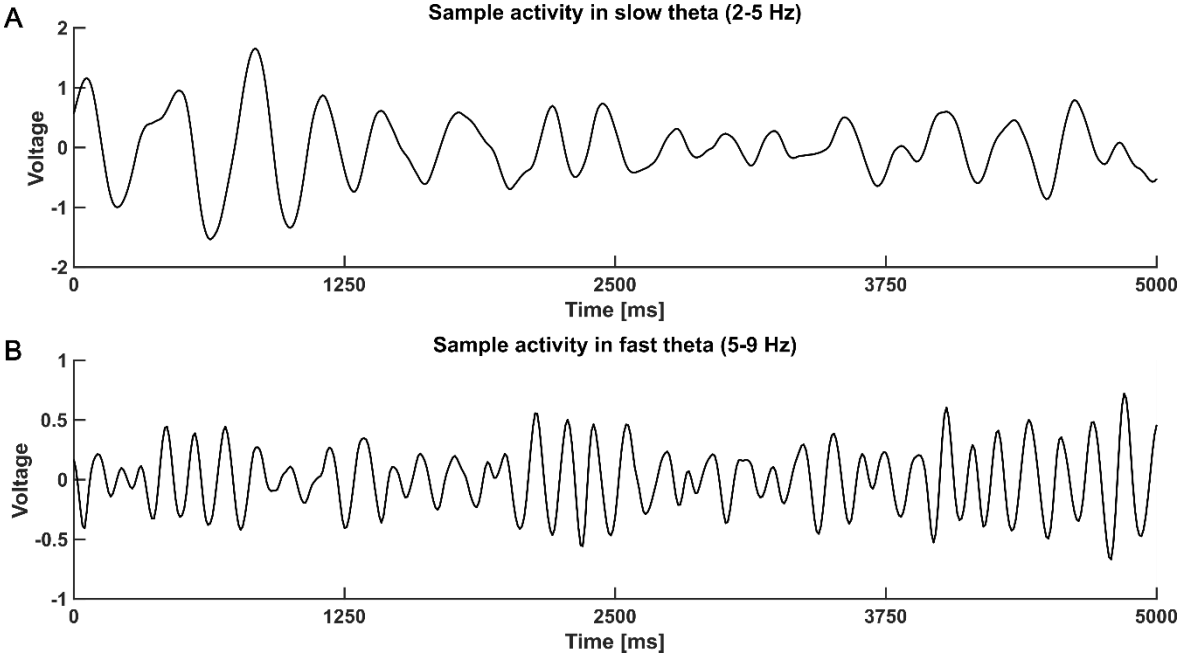


Figure 4.1. Five second data snippet showing activity in the slow (2-5 Hz; A) and fast (5-9 Hz; B) components.

Components were generated by taking a weighted average of the narrowband signal of all microwires within a bundle. The weighted average was calculated using a generalized eigendecomposition of the broadband and narrowband covariance matrices.

4.2.5 Spike-field coupling to slow and fast theta

We considered the spikes of neurons up to two seconds preceding the patient's response during the encoding and retrieval of later remembered episodes. Each neuron had to contain at least 11 spikes within the time of interest to be included for further analysis. We confined all spike-field analyses to spikes and LFPs that were recorded on the same Behnke-Fried electrode. We first wanted to estimate phase preference during encoding and retrieval independently. To do this we identified the complex value at the time of each spike. We subsequently normalized each complex value and averaged across spikes. For each neuron with spikes within the time of interest we computed the preferred phase by computing the angle of this average complex number. To estimate phase preference across neurons we performed a Rayleigh test. We next investigated whether there was a significant difference in the phase of the narrowband signal between spikes during encoding and retrieval for (i) Episode Specific Neurons in trials that were later reinstated (rESN), (ii) for Episode Specific neurons in trials that were later not reinstated (nESN) and (iii) all other neurons (SU). To this end, we computed the cosine similarity between the complex value of each spike at encoding with the complex value of each spike at retrieval. We then averaged these similarity values across spikes for each eligible neuron. We determined the statistical significance of

these difference scores using a one sample test for a mean angle of 0°, which we implemented using the function *circ_mtest* from the Circular Statistics Toolbox v1.21.0.0). However, if only few neurons are sensitive to the ongoing theta phase an encoding-retrieval phase offset in this small number might be overshadowed by other neurons whose activity is not theta modulated. To address this, we repeated the above phase difference analysis using only neurons whose spikes showed a significant coupling to the theta phase at encoding and retrieval, as evidenced by a Rayleigh test. We proceeded with this analysis only if there were at least 11 eligible neurons.

4.3 Results

We studied recordings from two different experiments (Experiment 1: 585 neurons and 1011 microwires in the hippocampus, 16 participants, 7 female; average age = 36.13 years, from 26-53 years; Experiment 2: 216 neurons and 339 microwires in the hippocampus, 14 participants, 7 female; average age = 33.86 years, from 19-58 years). Patients were implanted with stereotactic Behnke-Fried depth electrodes while completing a memory association task (see Figure 2.1). During the encoding phase of Experiment 1 patients were instructed to mentally create a vivid story consisting of an animal cue and two associate images (two faces, two places, or a face and a place). There was only one associate image in Experiment 2 and cue and associate could be either a face, a place, or an animal. Following a short distractor task where patients had to indicate whether a series of 15 numbers were odd or even the retrieval phase begun. During the retrieval phase the cue image was presented and the patient had to recall the associate image(s). Each episode was learned and retrieved only once, and the experiment was self-paced.

4.3.1 Periodic and aperiodic theta activity during correctly remembered and forgotten episodes

The power spectrum can be separated into periodic and aperiodic components. The periodic components reflect true oscillations, while the aperiodic component is also referred to as 1/f and is assumed to reflect general excitability (Gao et al., 2017). We separated periodic and aperiodic components in the microwire LFP using the FOOOF (Donoghue et al., 2020) implementation available in FieldTrip (Oostenveld et al., 2011) over a range of 1 Hz to 200 Hz and contrasted activity of later remembered with later forgotten episodes during encoding and retrieval (**Figure 4.2**). We found no significant differences in the aperiodic offset during

encoding in Experiment 1 (all $p > 0.54$) or Experiment 2 (all $p > 0.55$). However, during retrieval there was a significantly larger offset and steepness in the aperiodic signal for later forgotten episodes in both Experiment 1 (offset: $t_{offset}(341) = 3.13$, $mean_{remembered} = 2.23$ (*s.e.* = 0.047), $mean_{forgotten} = 2.25$ (*s.e.* = 0.050), $p_{offset} = 0.002$; steepness: $t_{tilt}(341) = 3.36$, $mean_{remembered} = 1.83$ (*s.e.* = 0.020), $mean_{forgotten} = 1.84$ (*s.e.* = 0.021), $p_{tilt} < 0.001$) and Experiment 2 (offset. $t_{offset}(114) = 3.00$, $mean_{remembered} = 2.04$ (*s.e.* = 0.084), $mean_{forgotten} = 2.08$ (*s.e.* = 0.088), $p_{offset} = 0.003$; steepness: $t_{tilt}(114) = 3.37$, $mean_{remembered} = 1.59$ (*s.e.* = 0.038), $mean_{forgotten} = 1.61$ (*s.e.* = 0.039), $p_{tilt} = 0.001$). We next compared the periodic theta activity between remembered and forgotten episodes. In Experiment 1, there was no difference in oscillatory slow or fast theta activity between the types of episodes during either encoding or retrieval (all p > 0.059). However, in Experiment 2 a difference in periodic fast theta activity during encoding emerged ($t(114) = 2.68$, $p = 0.008$; all other $p > 0.065$) where later forgotten episodes showed an increase in periodic power ($mean_{remembered} = 14.75$ (*s.e.* = 2.082); $mean_{forgotten} = 18.42$ (*s.e.* = 2.98).

	Contrast	Remembered vs forgotten episodes			
	Phase	Encoding (exp 1)	Encoding (exp 2)	Retrieval (exp 1)	Retrieval (exp 2)
Slow theta	Remembered episodes	$p = 0.603$	$p < 0.001$ $t(114) = 6.79$	$p = 0.025$	$p < 0.001$ $t(114) = 9.13$
	Forgotten episodes	$p = 0.076$	$p = 0.113$	$p = 0.009$ t(341) = 2.61	$p < 0.001$ $t(114) = 5.38$
Fast theta	Remembered episodes	$p < 0.001$ $t(365) = 7.41$	$p < 0.001$ $t(114) = 7.09$	$p = < 0.001$ $t(365) = 8.19$	$p < 0.001$ $t(114) = 6.61$
	Forgotten episodes	$p < 0.001$ $t(341) = 6.05$	$p < 0.001$ $t(114) = 6.18$	$p < 0.001$ $t(341) = 3.76$	$p < 0.001$ $t(114) = 6.51$

We found consistent evidence across experiments and experiment phase (i.e., encoding/retrieval) for periodic fast theta activity in remembered and forgotten episodes (all $p < 0.001$; see Table 4.1). In both experiments, forgotten trials contained significant slow theta activity during retrieval ($p < 0.009$) but not encoding ($p > 0.07$). Remembered episodes showed slow theta activity inconsistently across experiments. There was significant periodic activity during encoding and retrieval in Experiment 2 ($p < 0.001$), but not Experiment 1 ($p_{encoding} = 0.6$ and $p_{retrieval} = 0.025$).

To conclude, we observed an increased aperiodic offset and steepness for forgotten episodes compared to remembered episodes during retrieval, but not during encoding. There was no coherent difference in periodic slow or fast theta power between forgotten and remembered

episodes across experiments. We found reliable evidence for fast theta oscillations, whereas slow theta oscillations showed less clear results.

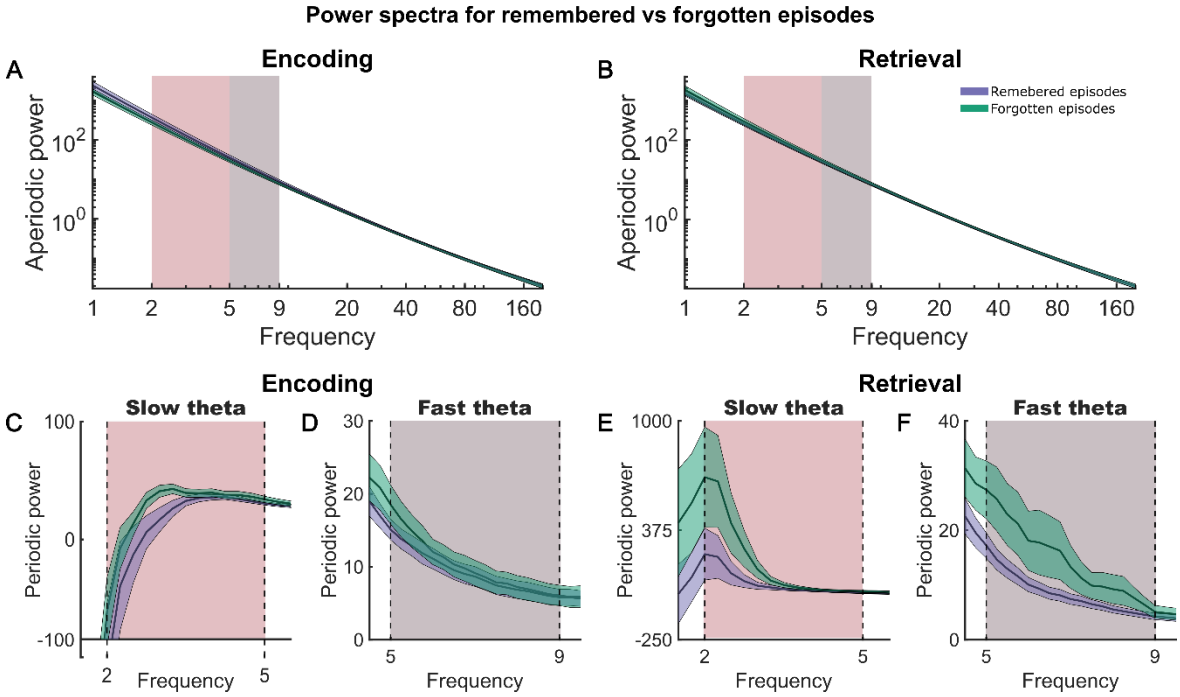


Figure 4.2. Aperiodic and oscillatory fast and slow theta activity during encoding and retrieval of remembered (purple) and forgotten (green) episodes. Green and purple shaded areas represented the SEM. Red and grey shaded areas represent the slow and fast theta frequency ranges respectively.

(A) Aperiodic power during encoding. Both axes are log-scaled. The x-axis shows frequencies from 1 to 200 Hz. The y-axis depicts the power at the respective frequency.

(B) Same as (A) but for retrieval

(C) Oscillatory activity in the slow theta range (2 Hz and 5 Hz) at encoding

(D) Oscillatory activity in the fast theta range (5 Hz and 9 Hz) at encoding

(E-F) Same as (C-D) but for retrieval.

4.3.2 Periodic and aperiodic theta activity during reinstated and non-reinstated episodes

We next contrasted periodic and aperiodic activity of reinstated against non-reinstated episodes on microwires that contained ESNs (Figure 4.3). We found no significant difference in the offset or steepness of the aperiodic component during encoding or retrieval in Experiment 1 (all $p > 0.3$) or Experiment 2 (all $p > 0.5$).

We then contrasted oscillatory activity in the slow and fast theta range between reinstated and non-reinstated episodes but found no significant differences during either encoding or

retrieval in Experiment 1 (all $p > 0.16$) or Experiment 2 (all $p > 0.09$). We found evidence for fast theta oscillations in reinstated and non-reinstated episodes during encoding and retrieval across both experiments ($p < 0.001$; see Table 4.2). There was no reliable pattern of slow theta oscillations across experiments when contrasting reinstated and non-reinstated episodes (see Table 4.2).

To conclude, despite finding evidence for the existence of theta oscillations, we did not find evidence for a difference in oscillatory power between reinstated and non-reinstated trials during encoding or retrieval. Likewise, there was no difference in aperiodic offset or steepness between later reinstated trials and non-reinstated trials during encoding or retrieval.

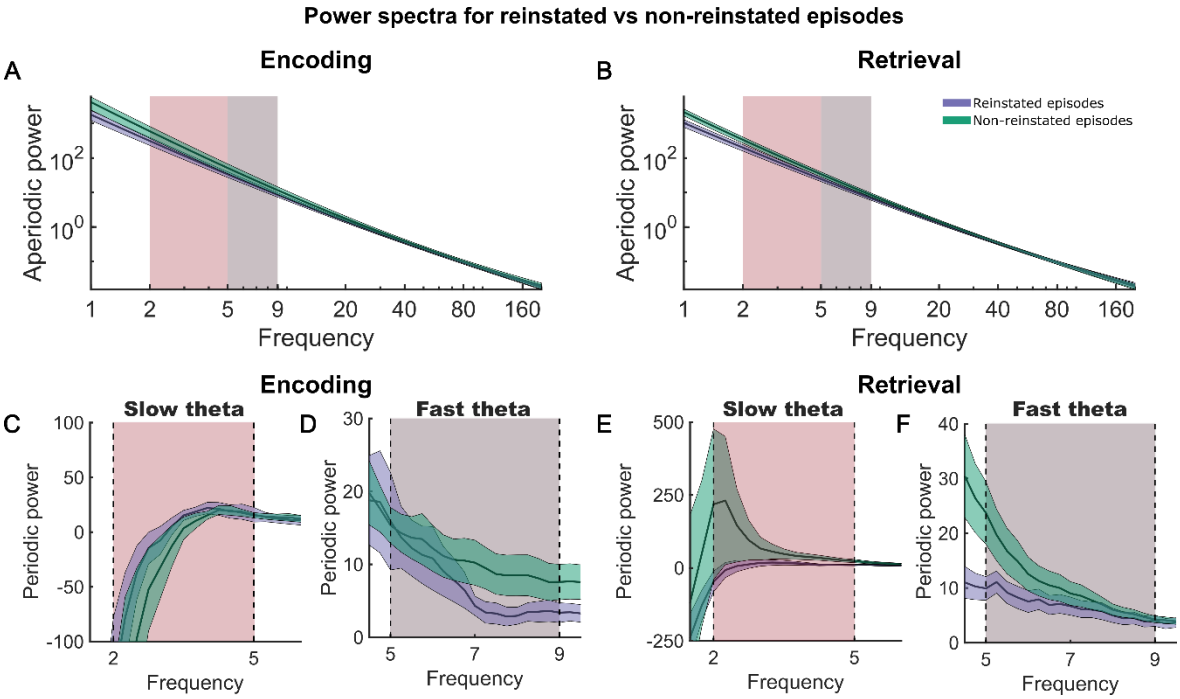


Figure 4.3. Aperiodic and oscillatory fast and slow theta activity during encoding and retrieval of reinstated (purple) and non-reinstated (green) episodes. Green and purple shaded areas represented the SEM. Red and grey shaded areas represent the slow and fast theta frequency ranges respectively.

(A) Aperiodic power during encoding. Both axes are log-scaled. The x-axis shows frequencies from 1 to 200 Hz. The y-axis depicts the power at the respective frequency.

(B) Same as (A) but for retrieval

(C) Oscillatory activity in the slow theta range (2 Hz and 5 Hz) at encoding

(D) Oscillatory activity in the fast theta range (5 Hz and 9 Hz) at encoding

(E-F) Same as (C-D) but for retrieval.

4.3.3 Single neuron firing to specific theta phases during memory encoding and retrieval

We next investigated whether single neuron firing would preferably occur within a specific theta phase during encoding and retrieval of episodic memories and whether there was a neuron specific phase offset between firing during the encoding and retrieval phases.

Based on previous literature no single theta frequency dominates the human hippocampus. Instead, there is a slower theta oscillation (2-5 Hz) and a faster theta oscillation (5-9 Hz) (Goyal et al., 2020; Kota et al., 2020).

We do not know which microwire best represents the dendritic input into a single neuron, so we computed theta components using a weighted average of all microwires within a microwire bundle. This was based on the generalized eigendecomposition of the narrowband theta covariance matrix and the broadband covariance matrix (see *4.2.4 Theta components and pre-processing*, p. 62). We distinguished three different categories of activity: spikes of ESN that occurred during reinstated trials (rESN; **Figure 4.4**), spikes of ESN during non-reinstated trials (nESN), and spikes of single units (SU; **Figure 4.5**). After excluding neurons with an insufficient number of spikes these analyses were based on $n_{rESN} = 36$, $n_{nESN} = 116$, and $n_{SU} = 380$ neurons in Experiment 1 and $n_{rESN} = 13$, $n_{nESN} = 34$, and $n_{SU} = 136$ neurons in Experiment 2. We first computed the preferred mean phase during encoding and retrieval for each neuron. To determine a general phase preference, we pooled this preferred phase value over all neurons within a category of neurons (rESN, nESN, SU) and used a Rayleigh test to determine statistically significant deviations from a uniform phase distribution. In Experiment 1, only the SU category showed a phase preference for the slow theta component during encoding ($\theta = 197.5°$, $p = 0.048$) and retrieval ($\theta = 181.9°$, $p = 0.004$). After adjusting for multiple comparisons for two tests (slow and fast theta) SU only showed a slow theta phase preference during retrieval ($p_{encoding\ adj.} = 0.096$; $p_{retrieval\ adj.} = 0.008$; Bonferroni corrected). Neither rESN nor nESN showed any slow or fast theta phase preference during encoding or retrieval (all $p > 0.28$).

In Experiment 2 the SU category showed a phase preference in the slow theta component during encoding ($\theta = 287.2°$, $p = 0.002$) but not during retrieval ($p = 0.633$; all other $p > 0.10$). There was a statistically significant phase preference of rESN for the slow theta component during retrieval ($\theta = 201.3°$, $p = 0.048$), however, after controlling for multiple comparisons (slow and fast theta), the effect was no longer significant ($p_{adj.} = 0.096$).

It is possible that despite an absence of phase preference during encoding and retrieval, neurons show a reliable offset between encoding and retrieval (a representative example of

a 10° offset with four neurons: encoding: 0°, 90°, 180°, 270°; retrieval: 10°, 100°, 190°, 280°). To determine if there was a significant theta phase difference between neurons firing at encoding and at retrieval, we computed the mean cosine similarity of the complex value for each neuron for all spikes during encoding with all spikes during retrieval. We determined the statistical significance of the encoding-retrieval phase offset separately for each neuron type (rESN, nESN, SU) using a one-sample test with a mean angle of 0° (i.e., no phase difference between encoding and retrieval). This one-sample test is the circular equivalent of a one-sample t-test with continuous data (we used the function *circ_mtest* from the Circular Statistics Toolbox v1.21.0.0). In Experiment 1 this approach yielded no significant encoding-retrieval phase differences for any category of neurons (rESN, nESN, SU) or theta components (slow, fast) (all $p > 0.26$). Likewise, no encoding-retrieval phase differences were found in Experiment 2 (all $p > 0.4$).

It is conceivable that theta activity modulates only some neurons. In this case a small proportion of theta-sensitive neurons might show a consistent phase difference between their firing at encoding and retrieval, but due to their small number this effect might be obscured. To circumvent this, we repeated the above phase difference analysis for neurons whose firing rate showed a phase coupling at encoding and retrieval using a Rayleigh test. Using this approach, we identified a significant phase offset between SU firing at encoding and retrieval in Experiment 1 ($\theta = 14°$, $p_{uncorrected} = 0.048$) that was no longer significant after correcting for multiple comparisons ($p_{corrected} = 0.096$; Experiment 1 all other $p > 0.39$; Experiment 2 all $p > 0.435$).

To conclude, we found a slow theta phase preference for SU during encoding in Experiment 2 and retrieval in Experiment 1. However, no neuron type (rESN, nESN, SU) showed a significant encoding-retrieval theta phase offset, which was also the case when limiting the theta phase offset analysis to neurons that showed a significant phase coupling at encoding and retrieval.

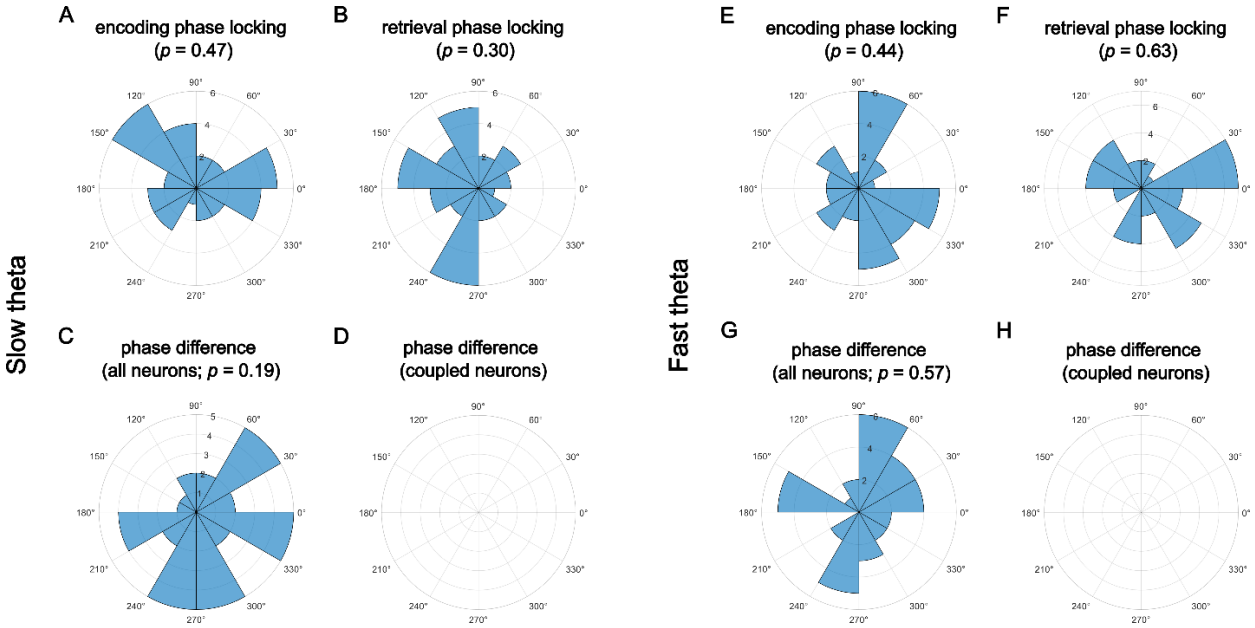


Figure 4.4. Polar histogram showing the phase offset between encoding and retrieval and the phase distribution during encoding and retrieval in ESNs during reinstated episodes.

(A) Preferred phase during encoding across all neurons for slow theta (2 Hz – 5 Hz)

(B) Preferred phase during retrieval across all neurons for slow theta (2 Hz – 5 Hz)

(C) Phase offset between encoding and retrieval across all neurons for slow theta (2 Hz – 5 Hz)

(D) Insufficient neurons (< 11) to analyse phase offset between encoding and retrieval in neurons showing significant theta coupling at encoding and at retrieval for slow theta (2 Hz – 5 Hz)

(E-H) Same as (A-D) but for fast theta (5 Hz – 9 Hz)

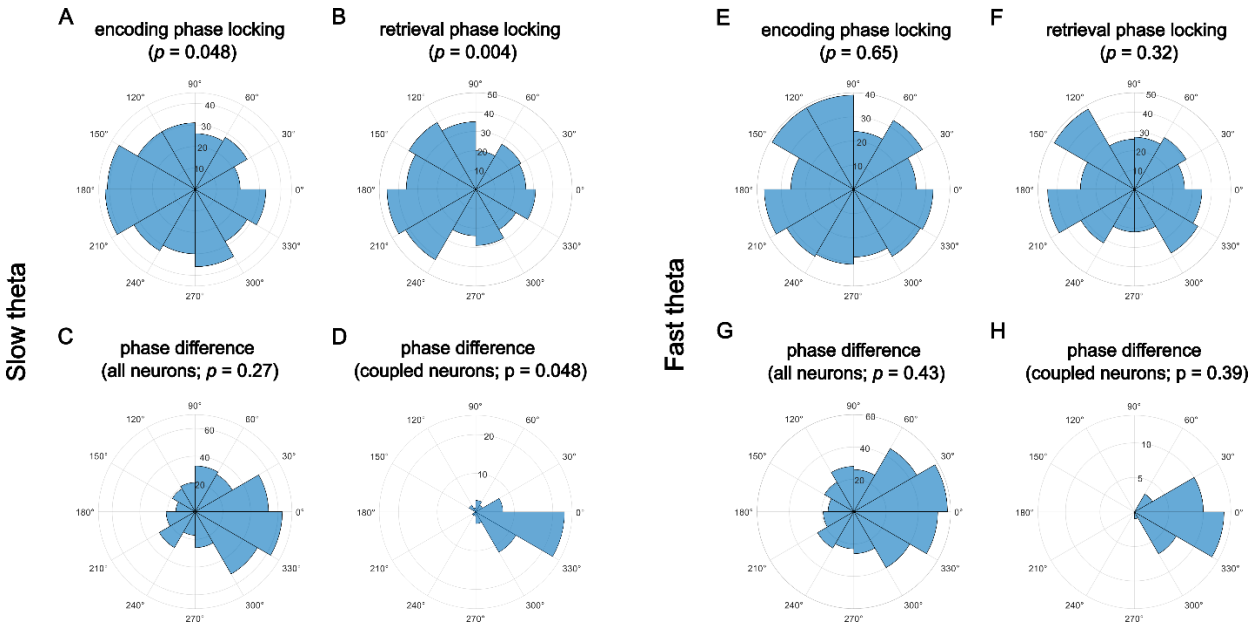


Figure 4.5. Polar histogram showing the phase offset between encoding and retrieval and the phase distribution during encoding and retrieval in non-ESN single neurons.

(A) Preferred phase during encoding across all neurons for slow theta (2 Hz – 5 Hz)

(B) Preferred phase during retrieval across all neurons for slow theta (2 Hz – 5 Hz)

(C) Phase offset between encoding and retrieval across all neurons for slow theta (2 Hz – 5 Hz)

(D) Phase offset between encoding and retrieval in neurons that showed a significant theta coupling at encoding and at retrieval for slow theta (2 Hz – 5 Hz)

(E-H) Same as (A-D) but for fast theta (5 Hz – 9 Hz)

Table 4.1. Overview of evidence for periodic fast and slow theta activity during encoding and retrieval of (later) remembered and (later) forgotten episodes in Experiment 1 and Experiment 2.

	Contrast	Remembered vs forgotten episodes			
	Phase	Encoding (exp 1)	Encoding (exp 2)	Retrieval (exp 1)	Retrieval (exp 2)
Slow theta	Remembered episodes	$p = 0.603$	$p < 0.001$ $t(114) = 6.79$	$p = 0.025$	$p < 0.001$ $t(114) = 9.13$
	Forgotten episodes	$p = 0.076$	$p = 0.113$	$p = 0.009$ $t(341) = 2.61$	$p < 0.001$ $t(114) = 5.38$
Fast theta	Remembered episodes	$p < 0.001$ $t(365) = 7.41$	$p < 0.001$ $t(114) = 7.09$	$p = < 0.001$ $t(365) = 8.19$	$p < 0.001$ $t(114) = 6.61$
	Forgotten episodes	$p < 0.001$ $t(341) = 6.05$	$p < 0.001$ $t(114) = 6.18$	$p < 0.001$ $t(341) = 3.76$	$p < 0.001$ $t(114) = 6.51$

Table 4.2. Overview of evidence for periodic fast and slow theta activity during encoding and retrieval of (later) reinstated or (later) non-reinstated episodes in Experiment 1 and Experiment 2.

	Contrast	Reinstated vs non-reinstated episodes			
	Phase	Encoding (exp 1)	Encoding (exp 2)	Retrieval (exp 1)	Retrieval (exp 2)
Slow theta	Reinstated episodes	$p = 0.318$	$p = 0.003$ $t(32) = 3.16$	$p = 0.495$	$p < 0.001$ $t(32) = 3.81$
	Non-reinstated episodes	$p = 0.196$	$p < 0.001$ $t(32) = 5.00$	$p = 0.1814$	$p < 0.001$ $t(32) = 3.48$
Fast theta	Reinstated episodes	$p < 0.001$ $t(122) = 3.42$	$p < 0.001$ $t(32) = 3.75$	$p < 0.001$ $t(122) = 4.65$	$p = 0.007$ $t(32) = 2.86$
	Non-reinstated episodes	$p < 0.001$ $t(122) = 3.82$	$p < 0.001$ $t(32) = 3.92$	$p < 0.001$ $t(122) = 4.52$	$p = 0.002$ $t(32) = 3.48$

4.4 Discussion

Episodic memories consist of various multimodal elements and are embedded in a distinct temporal and spatial context (Tulving, 1972; Tulving, 2002). The neurophysiological markers of episodic memory processing are still subject to debate, but a considerable body of literature exists that emphasizes the importance of theta oscillations for memory processing (Clouter et al., 2017; Griffiths et al., 2021; Hanslmayr et al., 2016; Hasselmo et al., 2002; Kerrén et al., 2018; Parish et al., 2021; Roux et al., 2022; ter Wal et al., 2021; Winson, 1978). We analysed the activity of single neurons relative to the ongoing theta activity in two independent intracranially recorded datasets that were collected using microelectrodes located in the human hippocampus while patients performed a memory association task.

In a recent review Herweg and colleagues (Herweg et al., 2020) suggested that memory processing is reflected in a steeper aperiodic component and an increase in periodic theta activity. Furthermore, studies have revealed that there is not one dominant theta frequency in the human hippocampus, but rather two distinct oscillations – a slow (2-5 Hz) and a fast (5-9 Hz) theta oscillation (Goyal et al., 2020; Kota et al., 2020). We first compared the aperiodic and periodic slow and fast theta components between remembered and forgotten episodes. In a second analysis we repeated the analysis but contrasted episodes during which the neural firing rate of ESNs is reinstated with episodes without neural firing reinstatement. In line with the hypothesis proposed by Herweg and colleagues (Herweg et al., 2020) we found a higher offset and 1/f tilt during retrieval of forgotten episodes. However, this aperiodic difference was absent during memory encoding, and we found no aperiodic differences between reinstated and non-reinstated episodes. We did not find any consistent differences in oscillatory slow and fast theta power for remembered vs forgotten episodes or reinstated vs non-reinstated episodes. We found periodic theta activity in both contrasts and during encoding and retrieval, although this evidence was more reliable in the fast theta band.

To conclude, evidence regarding the offset and steepness of the aperiodic component was inconclusive and we found no evidence of periodic theta power being involved in memory processing. There were no significant periodic or aperiodic differences between two categories of successful memory events (i.e., between reinstated and non-reinstated episodes). One possible reason may be that each successfully encoded memory is represented by an assembly of ESNs. As we recorded only from a small subgroup of the ones close to the microwires, the non-reinstated episodes would only differ insofar as we would not record their respective ESNs, because the LFP reflects a larger area than the area

in which spikes are recorded. However, there is a deeper problem with the argument presented by Herweg and colleagues (Herweg et al., 2020). They recommend contrasting the strength of two successful memories (e.g., how many contextual details are remembered when retrieving an episode). The idea behind that is that in both cases domain-general processes, reflected in the steepness of the aperiodic component, would be present and any differences would be driven by the memory strength. The problem with this is that it implicitly assumes that processes like task engagement, effort, perception and attention are binary. However, a more vividly remembered episode might have a shallower aperiodic component because the patient has paid more attention during the episode and not because of memory processing.

It should be noted that methods to separate periodic and aperiodic activity are far from perfect. Especially the large negative deflection in periodic activity e.g., at around 2 Hz in Figure 4.3C casts doubt on the validity of the aperiodic power estimation. Thus, oscillatory activity at the faster theta range may not reflect true periodicity but instead a poor 1/f fit. One might then argue that the lack of a consistent theta phase preference and no encoding-retrieval phase offset might be due to the absence of substantial periodic theta power in microwires. It is possible that macrowires instead integrate over larger areas and show more robust periodic theta activity (unless a bipolar reference is used; see Herweg et al., 2020). However, previous studies have shown spike-field-coupling in the theta range using microwires (Jacobs et al., 2007; Reddy, Zoefel et al., 2021; Roux et al., 2022; Rutishauser et al., 2010) and spikes can couple to the phase of aperiodic components (Bush & Burgess, 2020).

One influential theoretical model proposed that encoding and retrieval of memories occur in opposite phases of the theta oscillation thereby avoiding that encoding new information causes catastrophic interference of older memories (Hasselmo et al., 2002). We investigated how the firing activity of different previously identified neuron types relates to the phase of the ongoing theta oscillations during memory encoding and retrieval. We distinguished between spikes from ESNs during reinstated (rESN), non-reinstated episodes (nESN) and spikes from other single neurons (SU). Although we found some rudimentary evidence that SU show a slow theta (2-5 Hz) phase preference during encoding and retrieval, this finding was not consistent across the two experiments. Apart from that we did not detect any significant encoding or retrieval theta phase preference for neural firing across experiments. We also found no significant encoding-retrieval phase offset across all neurons, nor when limiting our analysis to neurons that showed significant theta phase coupling during encoding and retrieval.

These unexpected results could be due to various reasons. Many of our recorded neurons may not have been involved in active memory processing and thus did not show any modulation induced by memory encoding and retrieval. However, this does not explain our null findings for rESN, which are, by definition, coding for that specific episode. In this case, our results may be attributed to an insufficient number of eligible neurons or the two seconds preceding the patient's response may be a suboptimal time window for investigating spike-field coupling. Moreover, we did not differentiate between interneurons and pyramidal neurons, which are known to fire at different theta phases thus introducing more variance (Csicsvari et al., 1999).

Most neurons seem to maintain a preferred theta phase between encoding and retrieval. It is tempting to suggest that there is no theta phase preference during encoding and retrieval and that across the population of physiologically differently excitable neurons the entire theta cycle is covered leading to a uniform phase histogram at encoding and retrieval. However, we employed a frequentist approach when analyzing our data; thus, while we did not find compelling evidence to reject the null hypothesis (i.e., no theta phase difference between spikes at encoding and retrieval), this should not be interpreted as evidence for the null hypothesis (Dienes, 2014). To further investigate this, future studies should use a Bayesian framework and use a larger sample size.

To conclude the present chapter, in line with our hypothesis we find that forgotten when compared to remembered episodes have a higher aperiodic offset and a steeper gradient. Contrary to our hypotheses, we found no such pattern during memory encoding and no periodic theta increase for correctly remembered episodes. Likewise, we did not find evidence of neural firing in specific phases during encoding and retrieval, or a phase difference between encoding and retrieval in two independent datasets.

Chapter 5

General Discussion

Abstract

Throughout this thesis, I have investigated the neural underpinnings of episodic memory processing in the human hippocampus. In the second chapter, I approached this enigma on the level of single neurons. In the third chapter I examined it on the level of population activity as reflected in activity in the high frequency band. Finally, in the fourth chapter I examined the firing patterns of the previously identified neurons in relation to the phase of oscillatory theta activity. In this last chapter, I will summarize and integrate these findings, outline possible future experiments and discuss avenues for translational research.

5.1 Summary of findings

In the first chapter, I presented evidence of single neurons in the human hippocampus that reinstate their firing rate during the retrieval of specific episodes, both based on a rate code and a temporal code. These neurons, referred to as Episode Specific Neurons (ESNs), are distinct from neurons that are tuned to specific concepts (Concept Neurons; Mormann et al., 2008; Quian Quiroga et al., 2005) or reoccurring time points (Time Cells; Reddy, Zoefel et al., 2021; Umbach et al., 2020). Preliminary evidence indicates that these ESNs do not exist in the parahippocampus, although our coverage in that area is sparser than in the hippocampus. Additionally, initial evidence suggests that ESNs are likely to be excitatory pyramidal neurons.

In chapter 2, we extended these findings to the high frequency band in the local field potential. Although no consensus has been reached in the literature yet, it is generally agreed upon that an increase in high frequency power reflects an increase in local neural firing (Buzsáki et al., 2012; Manning et al., 2009; Nir et al., 2007; Ray et al., 2008). In line with this, we found a significant, albeit low, correlation between single neuron firing and high frequency power. In parallel to our earlier findings, we demonstrated that power in the high frequency band (40-200 Hz) was reinstated for particular episodes in a significant number of microwires. This finding was limited to later remembered episodes and did not emerge for later forgotten episodes. Although we did find stimulus dependent high frequency power (HFP) modulations akin to the firing rate increases in Concept Neurons, these HFP changes were not responsible for the HFP memory reinstatements in ESWs. Unexpectedly, the relative power increases in reinstated episodes extended past our frequency range of interest (until ~10 Hz during memory encoding and ~15 Hz during memory retrieval). Future studies should differentiate whether this finding is driven by a power offset, an aperiodic 1/f shift, or oscillatory components.

Ample research points towards a central role of theta oscillations in episodic memory processing. Recent work has shown that hippocampal theta oscillations in humans is divided between a slow (2-5 Hz) and fast (5-9 Hz) theta oscillation. However, the exact role of theta remains elusive. Herweg and colleagues (Herweg et al., 2020) proposed that successful memory is reflected in a power shift towards higher frequencies and a circumscribed narrow-band periodic theta increase. In the third chapter we tested this hypothesis in two intracranial datasets but did not find consistent evidence to support it. Based on influential theoretical work by Hasselmo and colleagues (Hasselmo et al., 2002), we expected non-ESN single neurons and ESNs to lock onto different phases of theta. Contrary to our hypothesis, we were

unable to identify a consistent phase preference during encoding or retrieval of episodic memories across the two independent datasets. We also did not find evidence for a theta phase offset between encoding and retrieval. Indeed, many neurons might not be involved in processing a given memory, which provides a possible reason for our findings. However, this does not apply to ESNs which, by definition, code for a specific memory as reflected in their increased firing rates for that memory. The absence of a theta phase effect in this case may be attributed to the low number of ESNs leading to insufficient power to detect an effect. Although a complete absence of theta phase preference goes against previous findings, neurons reportedly lock to a large range of theta phases (Jacobs et al., 2007; Reddy, Self et al., 2021; Roux et al., 2022; Rutishauser et al., 2010).

To conclude, we found a single neuron basis of memory processing and extended these findings to activity in a greater population of neurons reflected in the local field potential. While many exciting open questions remain, we hope to have laid a foundation for future work. In the following text we sketch some of these questions.

5.2 ESNs and Index Neurons

Although our research, which culminated in compelling evidence for ESNs, was inspired by what Teyler and DiScenna (Teyler & DiScenna, 1986) called Index Neurons, we did not call them such. This is because there are features ascribed to Index Neurons that we cannot test using the two available datasets. Upon presentation of a partial input present at memory encoding the Index Neuron assembly in CA3 is pattern completed. We cannot test this, because we always use the same memory cue for encoding and retrieval in both of our experiments. Nevertheless, we predict that using varying memory cues from the same memory should reinstate the same ESNs. Having access to multiple neurons that are allocated to the same episode would create the opportunity to ask more nuanced questions. If a cue does not initiate memory retrieval, do some neurons reinstate their firing activity, but the activation is not sufficient to pattern complete the entire assembly? Is there a relationship between the strength of assembly reinstatement and the detail of episodic retrieval? Is there a specific reactivation order based on the memory cue, which would indicate that within an assembly, individual neurons are responsible for specific features within the episode? It is conceivable that depending on the hippocampal subfield these answers differ. As we lack sufficient coverage to record multiple neurons of one assembly that codes an episode, we cannot currently investigate these questions. Conversely, pattern separation should allow the distinction of highly similar, but different episodes by assigning

them to different Index Neuron ensembles. We are unable to verify this using the current experiments because the images used in each episode do not overlap.

Another important question is the stability of the ESN code over time and multiple retrievals. We expect ESNs to generally reinstate their firing pattern on all subsequent retrievals. However, some ESNs may drop out of the assembly that is allocated to a given episode, which would lead to some variance. Patients in our studies retrieved every episode only once, so we cannot investigate this question. We are currently running an experiment where each episode is retrieved multiple times but will not have a complete dataset in the foreseeable future. An interesting, related endeavour is exploring the stability of ESNs over time. Are memory traces, as evidenced by ESN firing and HFP increases, systematically reactivated during the periods of (extended) consolidation? In this context, sleep is of particular interest as numerous studies suggest it has a role in memory consolidation by reactivating previous experiences (Born et al., 2006; Kolibius et al., 2020; Schreiner et al., 2021). Do ESNs reliably reinstate a memory days and weeks after memory encoding? Note, that it is a separate question whether the hippocampus in general stays involved in older memories (Nadel & Moscovitch, 1997) or not (McClelland et al., 1995; Squire & Alvarez, 1995). This is because it is conceivable that the original memory trace is transformed during the consolidation period between encoding and retrieval, meaning the initially allocated neurons are replaced by other neurons or pruned. It is easy in this case to erroneously infer that the hippocampus becomes redundant in retrieving a distant episodic memory when in reality the hippocampal memory trace persists in an altered form (i.e., consisting of fewer or other neurons). In light of this, being able to record more single neurons or even multiple ESNs that reinstate the same episode would be especially insightful. In absence of this possibility, our findings that episode reinstatement concurs with HFP increases could be used to investigate the stability of a memory code and the continuous involvement of the hippocampus in episodic memories. As a proxy of much broader neural firing, HFP is likely robust to a drop out of some of the initially allocated neurons.

5.3 Information flow between hippocampus and neocortex

A central part of the Indexing Theory is that ongoing cortical activation is bound by neurons in the hippocampus which project back and reactivate the initial cortical pattern during successful retrieval (Teyler & DiScenna, 1986; Teyler & Rudy, 2007). There are several hurdles in showing this empirically. One is a low spatial coverage of implanted electrodes, which, on top of that, is different for each patient as not all patients have electrodes in the

same part of the hippocampus or neocortex. Apart from that, the number of ESNs per patient is low. These concerns should not deter the curious reader, but merely caution to the difficulty of the task. A persisting problem has been discerning periods of memory reinstatement from background activity. One approach to this problem might be combining high frequency activity with concurrent ESN firing, which could serve as a more accurate indicator of memory reinstatement. A second approach is to correlate the instantaneous firing rate of ESNs or HFP with the output of a classifier (e.g., a linear discriminant analysis), which represents a retrieved memory in the neocortex. This is preferable over computing the classifier evidence at each individual spike which would assume that each spike leads to memory reinstatement. Especially in neurons that spike frequently each spike might only reflect baseline firing. Furthermore, focusing on each individual spike effectively smoothes the evidence for memory retrieval which makes it difficult to compare ESNs with different firing rates. A larger increase in instantaneous firing rate would also be easier to discern from baseline firing using a thresholding procedure and could be used as a marker for memory reinstatement. A third way to detect memory reinstatement would be to apply a time-based causality measure (e.g., Granger causality, Granger, 1969; Phase Slope Index, Nolte et al., 2008) on shorter data segments during the retrieval phase, when information is thought to flow from the hippocampus to the neocortex (Lifanov et al., 2022; Linde-Domingo et al., 2019). Segments in which activity in one area predicts the activity in the other are likely timepoints at which memory retrieval occurs. A fourth way would be to identify memory reinstatement in the neocortex through a classifier and reverse engineer the hippocampal activity pattern that induced it (e.g., looking at the neural activity one second prior). Finally, a fifth way would be applying a classifier to the hippocampal and cortical recordings separately and cross-correlate the two outputs or apply a causality test. In this case one would not focus on the transfer of the signal as raw data, but rather the transfer of evidence for content that represents that memory. To complicate things further one may wish to use the instantaneous firing rate and/or data within the high frequency band instead of the unfiltered raw micro-/macrowire recordings as inputs to these analyses. To conclude, only future studies using other experimental designs and/or more advanced hardware will be able to ascertain whether ESNs are indeed Index Neurons. Until then, we must be satisfied to see a reinstatement of neural firing as an indicator of memory processing.

5.4 On the origin of Concept Neurons

Recent times have seen an explosion in neuron types. We find cells that code for spatial locations, such as place cells (O'Keefe & Dostrovsky, 1971), grid cells (Hafting et al., 2005), head-direction cells (Taube et al., 1990), or egocentric bearing cells (Kunz et al., 2021). In humans, neurons that code for specific concepts, so called Concept Neurons have been found consistently (Mormann et al., 2008; Quian Quiroga et al., 2005). Recent additions include novelty cells and familiarity cells (Rutishauser et al., 2006; Rutishauser et al., 2008; Rutishauser et al., 2015) as well border and event cells (Zheng et al., 2022). In this work we introduced Episode Specific Neurons and added to this veritable embarras de richesse. I do not believe that each of these neurons represent a physiologically distinct neuron type that earns its label through a separate coding mechanism. I would like to therefore muse on the open question how CN initially develop their tuning. One possibility is that over repeated reconsolidation CNs evolve from ESNs. Imagine you meet your best friend in a coffee shop. This coffee shop episode will initially be represented by an assembly of ESNs. A few days later you meet with the same friend in a park and you remember the last time you met in the coffee shop. This reactivates the ESNs coding for the coffee shop episode. Engram literature suggests that recently active and more excitable neurons are preferentially bound to a new episode (Frankland & Josselyn, 2015; Josselyn, 2010). This makes it likely that some of the ESNs that coded the coffee episode now also code the park episode. The common element between those two episodes is your best friend. It is conceivable that over many such episodes a proportion of the ESNs that initially coded the coffee shop episode would become "semanticized", i.e., develop a tuning for your best friend. A Concept Neuron is born. In line with the hypothesis that Concept Neurons have been shown to change their tuning based on statistical regularities in the environment (Ison et al., 2015). In this way ESNs can be likened to variables in a computer program to which arbitrary information is bound. In the case of episodic memory, this arbitrary information would be the complete set of features that make up an episode. Alternatively, CN and ESNs might be located in different hippocampal subfields.

5.5 More advanced electrodes

The yield of neurons using the currently available electrodes is about a dozen per microwire bundle. In comparison, electrodes developed more recently yield hundreds of well localized neurons. Virtually all analyses described in this thesis would benefit from more neurons and

I have mentioned this throughout the manuscript when appropriate. Not in all cases can this difference in neuron yield be made up by recording more participants or sessions. For instance, recording hundreds of neurons in one patient would enable the analysis of between-cell interactions, such as the detection of assemblies of Episode Specific Neurons. Something that is not feasible when recording few neurons. Beyond this, between area interactions could be investigated (Durand et al., 2022). One concrete example would be a reinstatement of neocortical neurons that represent a memory and their interaction with hippocampal ESNs. During encoding one would assume that the neocortical neurons drive ESNs, while during retrieval the neocortex lags behind the hippocampus (Lifanov et al., 2022; Linde-Domingo et al., 2019).

5.6 Microwire stimulation

As I am writing this Neuralynx (Neuralynx Inc, USA) is seeking CE and FDA approval for microwire stimulation in patients (personal communication). If successful, microwire stimulation could provide causal evidence for an ESN based memory code. If ESNs are allocated based on excitability, as predicted by experiments in rodents, stimulating the neurons in the vicinity of a microwire should increase the probability that they are allocated to an episode. In other words, through electric stimulation hippocampal neurons may be galvanized - at the push of a button - into coding for a particular episode. Apart from shedding light on the mechanism by which ESNs are allocated this would also greatly increase the yield of ESNs per patient. Using this method, one could test the hypothesis that CN develop from ESNs by stimulating on a microwire during multiple episodes that share a common element (e.g., Jennifer Aniston in Pisa, Jennifer Aniston in Paris, …). If the stimulation causes neurons to be co-allocated to these episodes it is conceivable that some of the “tagged” neurons exhibit neural firing akin to Concept Neurons tuned to Jennifer Aniston (as she is the common element in these episodes). Likewise reactivating the previously assigned neurons through electrical stimulation should increase the probability the memory is retrieved or even induce memory retrieval in the absence of a retrieval cue. Conversely, depending on the effect of the stimulation protocol, memory retrieval may be prohibited.

5.7 Translational applications

In our everyday lives, we constantly generate and retrieve episodic memories - a marvellous ability we usually don't think about. This makes it even more painful and disruptive when memory processing becomes dysfunctional. Alzheimer's disease is a progressive degenerative disorder that causes atrophy of neurons resulting in one of its most recognizable symptoms - memory loss (World Health Organization, 2019, 2022). Alzheimer's is the most common form of dementia and was the 7th leading cause of death in the USA in 2020 (Murphy et al., 2021). Worldwide around 55 million people have dementia, a number which is expected to rise to 78 million by 2030 (World Health Organization, 2022). I am hopeful that our research can contribute to alleviate some of the burden that comes with Alzheimer's, although a cure requires a more targeted understanding of the disease. However, it might prove fruitful to electrically stimulate parts of the hippocampus during the early phases of the disease. This could help with the recruitment of ESNs to new memories or reinstate existing memories and thereby temporarily increase their quality of life. Early advances in developing a hippocampal neural prosthetic have shown great promise (Hampson et al., 2018).

Another disorder characterized by dysfunctional memories is the Post Traumatic Stress Disorder (PTSD). Its symptoms include vivid intrusive memories or flashbacks that can be triggered by cues reminiscent of the traumatising event. Re-experiencing the traumatic memory can cause intense physical and strong emotional reactions (World Health Organization, 2019). Although PTSD patients often suffer from deficits in declarative memory (Samuelson, 2011), a core issue is the pathological remembering of a disturbing event. Studies in rodents suggest that the size of the hippocampal engram does not increase with the intensity of the memory (Choi et al., 2018; Rao-Ruiz et al., 2019). This raises the question whether it may be possible to interact with the hippocampal memory trace that represents a traumatic experience. Is it possible to downregulate the engram on demand, or even ablate it? The latter raises an important ethical concern - does erasing an entire experience also abolish the consequences of the experience or does it simply get rid of the conscious perception leaving the visceral reaction intact? Of course, the idea of simply "switching off" a memory as a treatment simplifies the underlying syndrome of PTSD. Nonetheless, the idea is that through basic research such as the work presented here we gain a better understanding of how memories are processed in humans and that this insight helps translational applications further down the road.

Appendix: Supplementary Material for Chapter 1

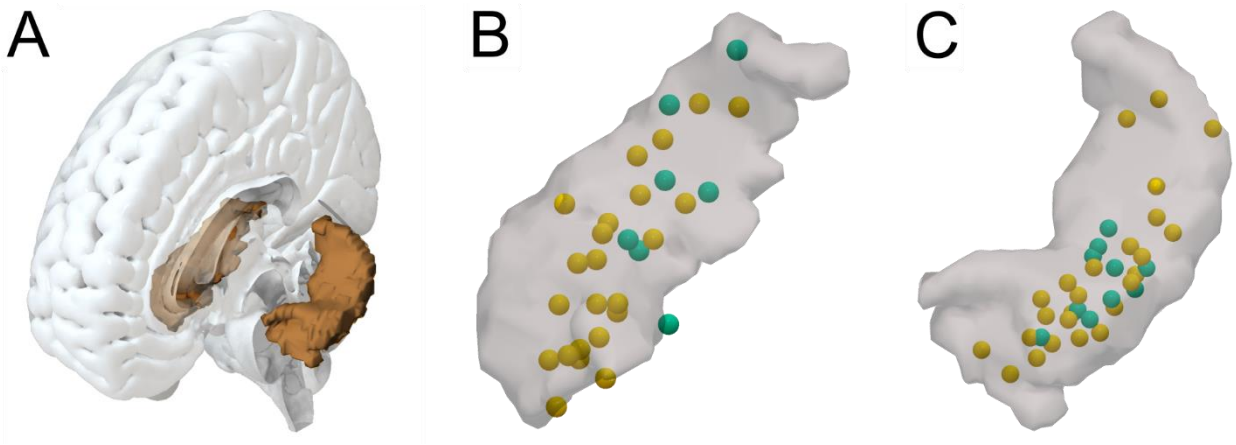


Figure A.1. Visualisation of the hippocampus and electrode positions for Experiment 1 and Experiment 2.

(A) Outline of the hippocampus within a whole brain mesh.

(B). Normalized right hippocampus. The yellow spheres represent the estimated position of microwire bundles that contain ESNs. The green spheres represent the estimated position of microwire bundles that do not contain ESNs. Only bundles where single-unit activity was recorded are shown.

(C) Same as (B), but for the left hippocampus.

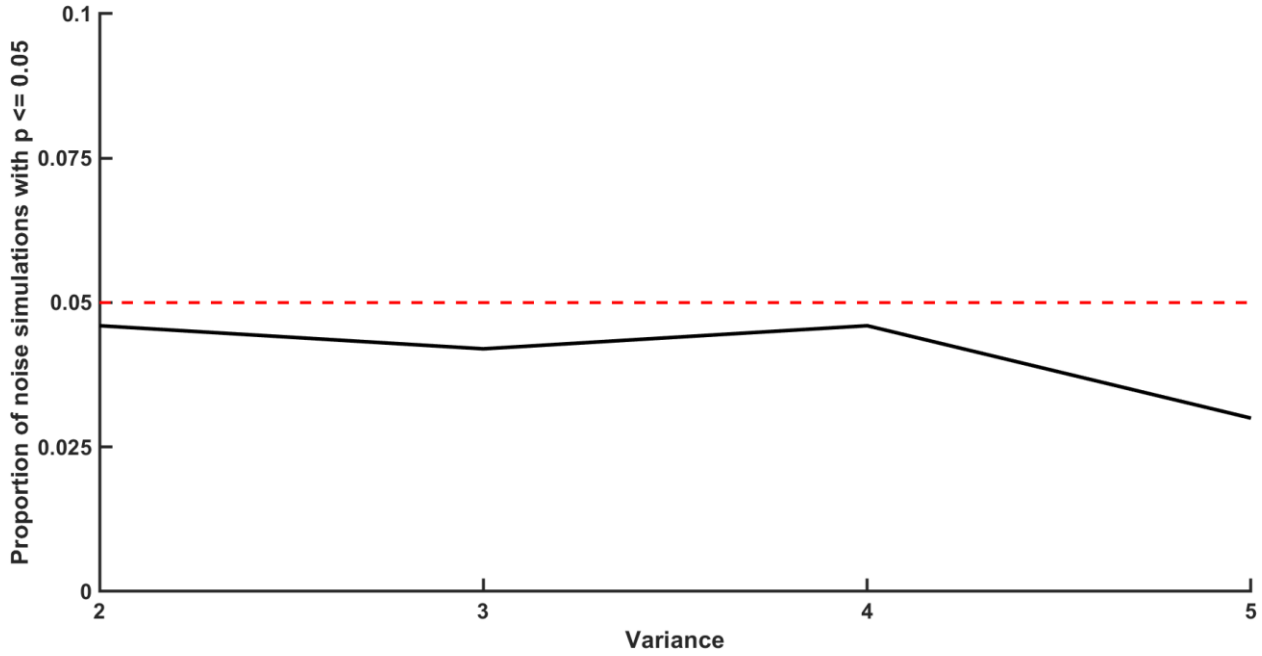
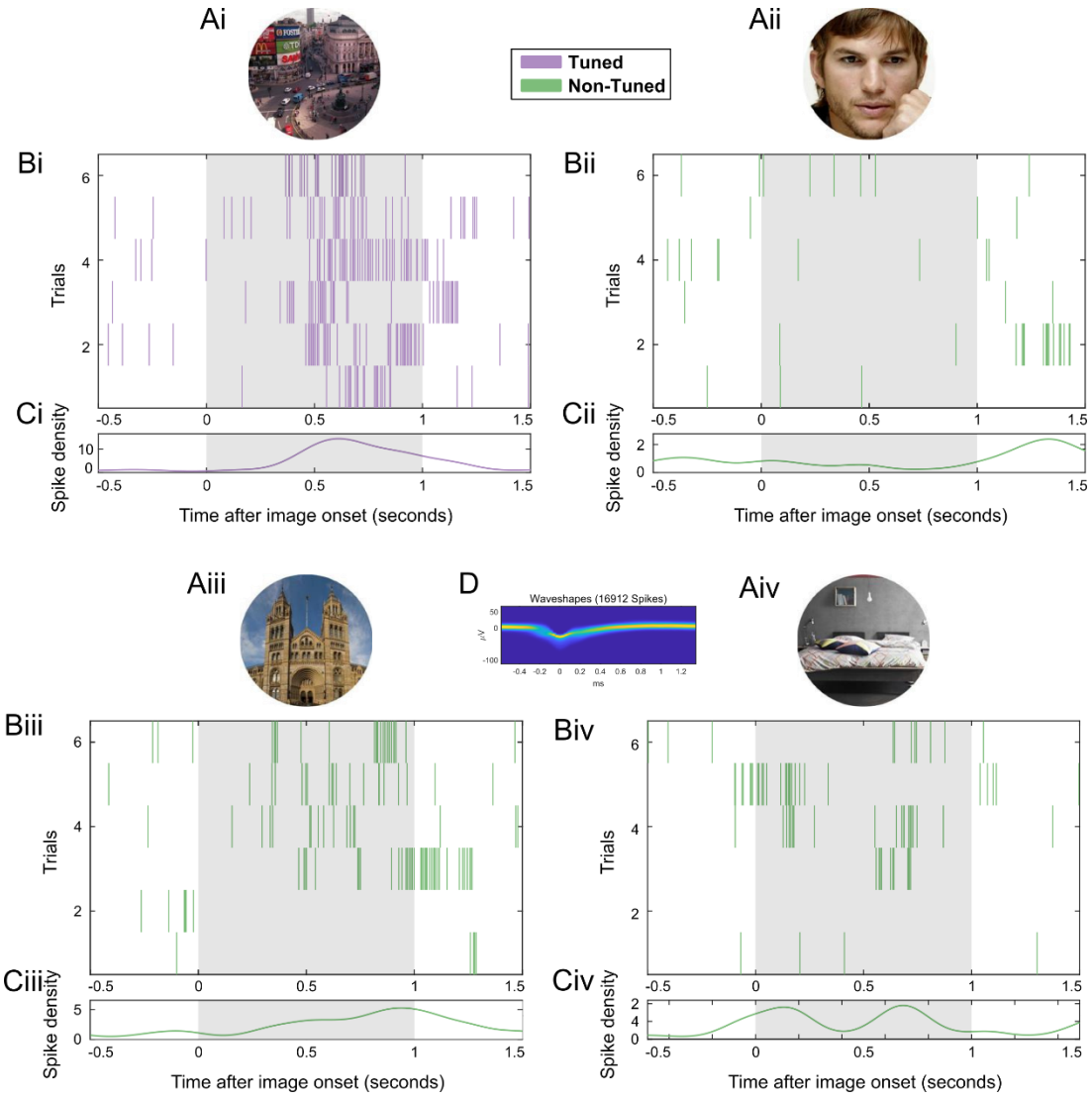


Figure A.2. Simulation of ESN identification.

We created a simulation using random pseudo spike rates to determine whether our ESN analysis pipeline contains a bias towards significant results over multiple levels of variance (x-Axis). For each level of variance, we repeated this step 1,000 times and calculated the proportion of iterations that yield a significant result (y-axis; p <= 0.05). The dotted red line represents the 5%-level, and the straight black line represents the results from the stimulation.


Ai
Aii
Tuned
Non-Tuned
Bi
Bii
Trials
Ci
Cii
Spike density
Time after image onset (seconds)
Aiii
D
Waveshapes (16912 Spikes)
Aiv
Biii
Biv
Ciii
Civ

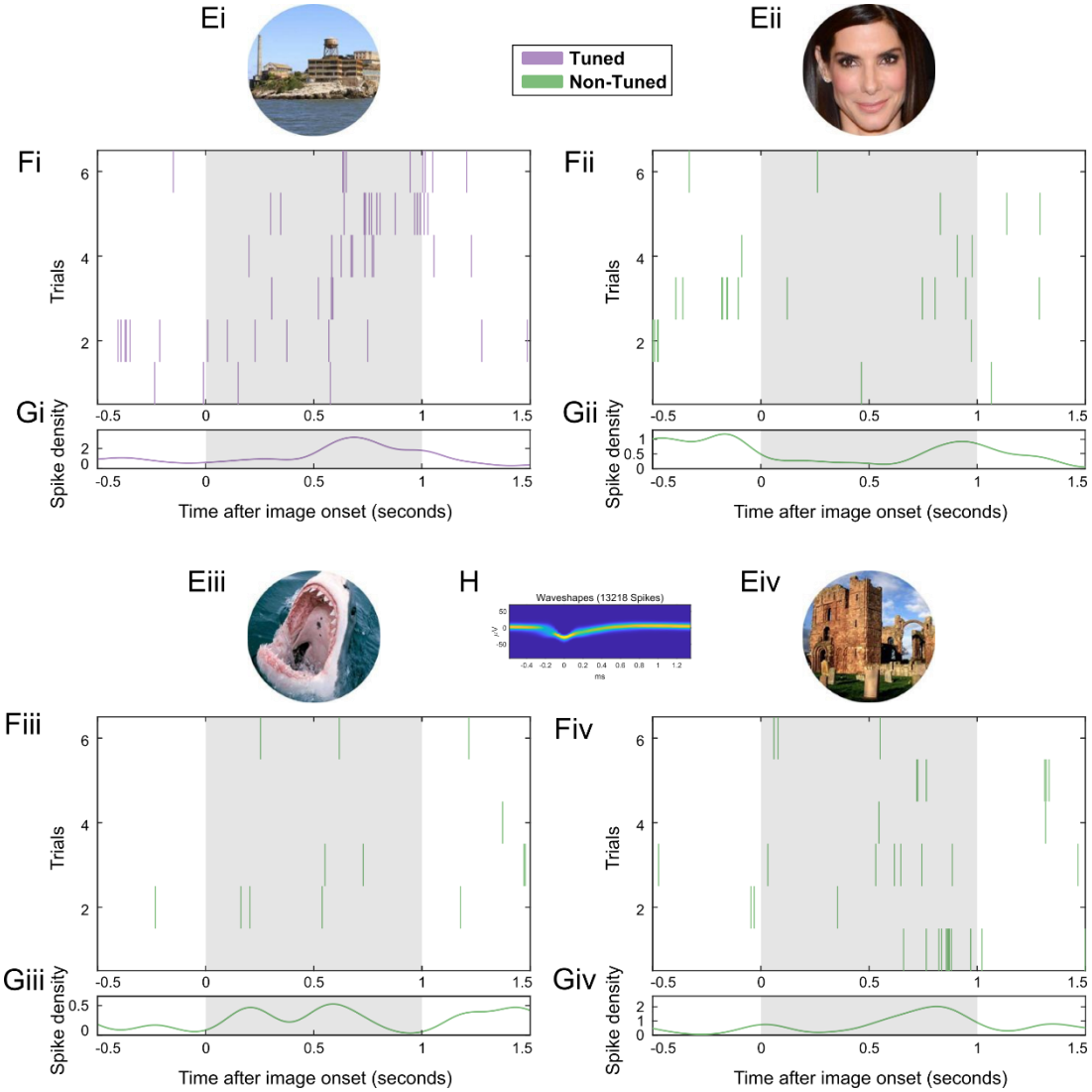


Figure A.3. Firing patterns for two example putative Concept Neurons that were identified using a visual tuning task in Experiment 2.

(A) Each during the memory task previously shown image (either an animal, a face or a place) is shown six times during the visual tuning task.

(B) Spike raster plot. Each line indicates a spike. On the x-axis is time (locked to image presentation) and on the y-axis are the six trials during which the above image is shown on the screen. Color-coded in purple for tuned images and green for non-tuned images. The grey area indicates the activation period that is considered for identifying Concept Neurons.

(C) Spike density plot (mean instantaneous firing rate over all six trials).

(D) 2D histogram of the waveshape of that particular unit (Niediek et al., 2016).

(E-H) same as (A-D) but for a different example ESN.

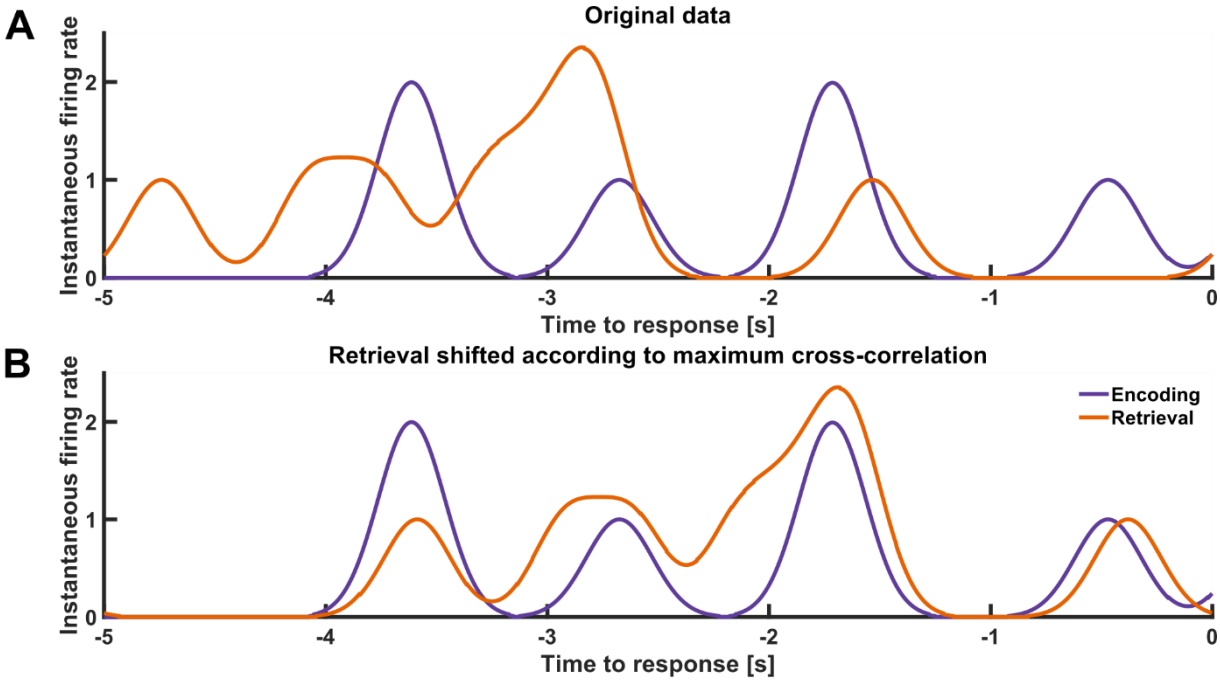


Figure A.4. Example of a trial reinstated by an example temporal Episode Specific Neuron (tESN).

(A) The original instantaneous firing rate during encoding (blue) and retrieval (orange) 5000ms before the response.

(B) Same as (A), but with the retrieval firing rate shifted according to the peak in the cross-correlation.

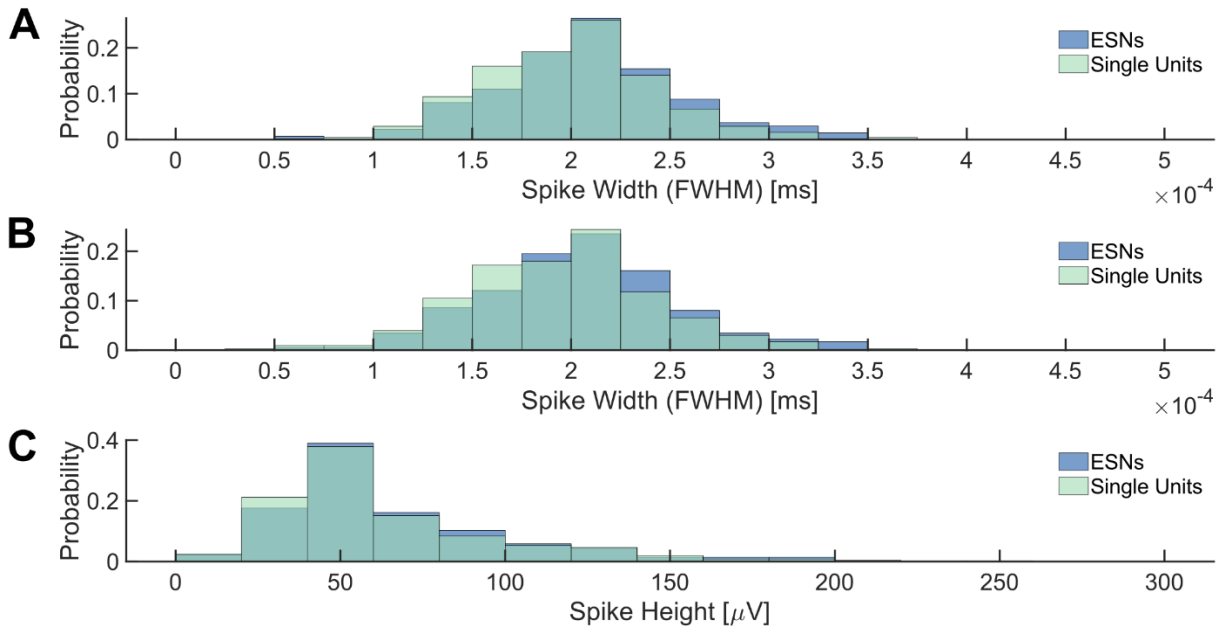


Figure A.5.

(A) Distribution of spike widths for neurons of the first experiment.

(B) Same as (A) but for neurons of the first and second experiment combined.

(C) distribution of spike height of ESNs (purple) and single units (green).

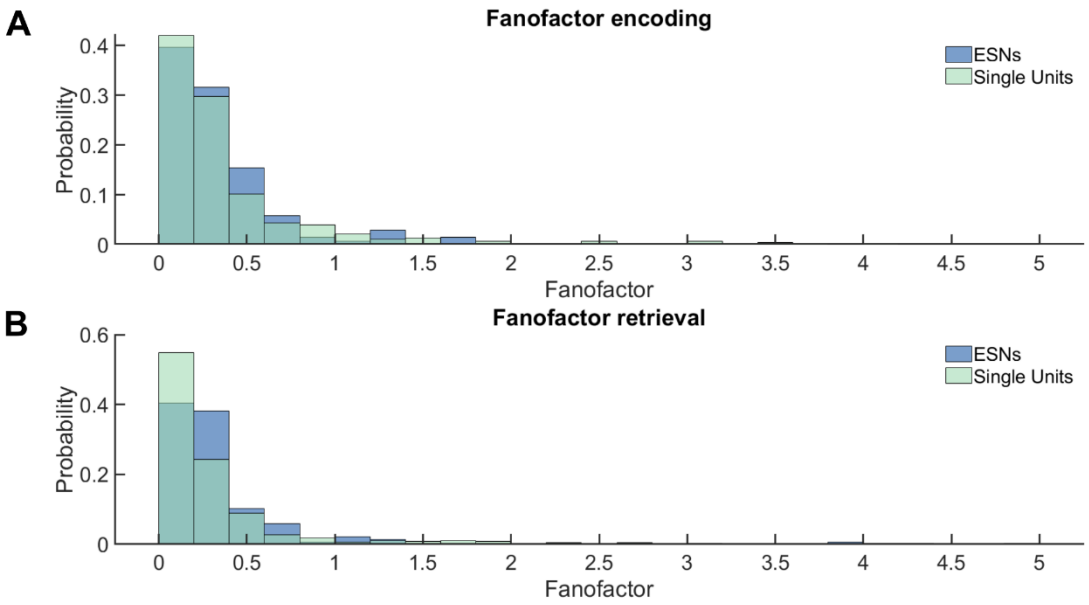


Figure A.6.

(A) Distribution of Fano factors of ESNs (purple) and single units (green) during encoding and during

(B) retrieval.

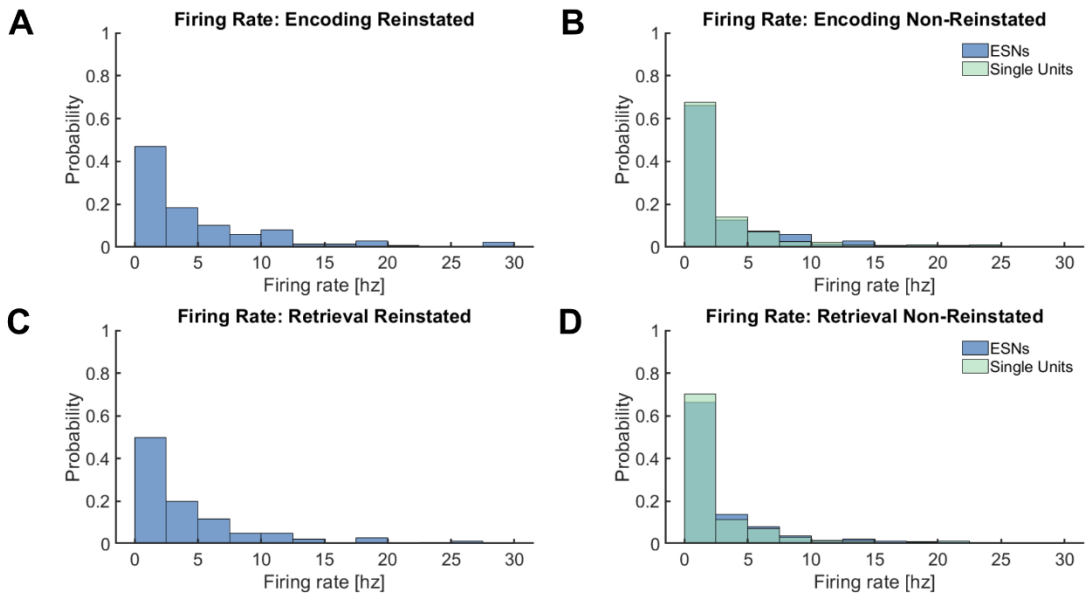


Figure A.7.

Firing rate in hertz of ESNs (purple) during

(A) encoding and

(B) retrieval of reinstated episodes.

Firing rate of ESNs and other single units (green) during

(C) encoding and

(D) retrieval of non-reinstated episodes.

Table A.1. Overview of electrode implantation and memory performance in Experiment 1.

Patient ID	Number of sessions	Number of bundles in hippocampus	Trial number[a]	Hits[a]	Hipp. bundles with SUs[a,b]	Number of hipp. SUs[a,b]	Number of ESNs[a,c]
0002	7	6	49.4 (0.4)	43.6 (1.3)	2.6 (0.3)	12.3 (2.9)	2.7 (0.64)
0004	3	3	49 (0)	33.7 (0.7)	2 (0)	7.3 (2)	1 (0.58)
0005	4	6	49 (0)	41.5 (2.7)	3.3 (0.3)	10.3 (1.5)	3.8 (0.63)
0007	3	4	94.3 (1.7)	86.7 (4.9)	4 (0)	16.7 (4.9)	8.7 (2.6)
0008	4	4	68.8 (2.3)	39.3 (3.9)	1.8 (0.3)	8.3 (1.6)	1.5 (0.29)
0009	3	5	84.3 (7.3)	56.3 (7.3)	4 (0)	17 (1.2)	2.7 (1.2)
0012	4	6	53.8 (8.1)	32.5 (6.9)	4 (0.4)	21.5 (4.1)	4.5 (0.87)
0013	6	5	76.2 (6.3)	45.3 (8.6)	2.8 (0.3)	17.2 (1.6)	3.7 (1.3)
1003	4	1	52.5 (2.4)	46.3 (4.9)	1 (0)	6.3 (0.5)	1.5 (0.87)
1004	2	2	84.5 (11.5)	75 (11)	1 (0)	1 (0)	1 (0)
1005	4	2	73 (13.9)	33.3 (7)	1.5 (0.3)	3.5 (1)	0.5 (0.29)
1007	5	1	49.6 (8.5)	31 (6.3)	1 (0)	4.4 (0.5)	0.8 (0.37)
1008	3	1	85 (9.5)	22.3 (6.8)	1 (0)	1.7 (0.7)	0 (0)
1009	2	1	82.5 (13.5)	67.5 (12.5)	1 (0)	2.5 (0.5)	0 (0)
1011	2	2	51 (10)	38 (5)	1.5 (0.5)	8.5 (0.5)	1.5 (1.5)
1012	3	2	39.7 (11.6)	21 (4.9)	1 (0)	7.7 (0.7)	0.67 (0.67)

[a] Each number stands for the mean over all experimental sessions with the standard error across sessions in brackets. [b] SUs: Single Units (including ESNs).[c] ESNs: Episode Specific Neurons

Table A.2. Overview of electrode implantation and memory performance in Experiment 2.

Patient ID	Number of sessions	Number of bundles in hippocampus with SUs	Trial number[a]	Hits[a]	Number of hipp. SUs[a,b]	Number of ESNs[a,c]
1013	1	1	49 (0)	37 (0)	16 (0)	3 (0)
1014	1	1	70 (0)	28 (0)	1 (0)	0 (0)
1015	3	1.7	62 (0.6)	28 (1.8)	9 (0.6)	2.3 (0.3)
1016	3	2	46.7 (4.1)	31.3 (3.8)	17.3 (1.5)	3.3 (0.3)
1017	2	1.5	62 (2)	38.5 (0.5)	2 (1)	0 (0)
1018	2	3	56 (9)	31 (5)	16 (1)	1.5 (0.5)
1019	1	2	64 (0)	49 (0)	12 (0)	5 (0)
1020	1	2	51 (0)	27 (0)	22 (0)	2 (0)
1021	1	2	54 (0)	38 (0)	15 (0)	1 (0)
1022	1	2	62 (0)	28 (0)	10 (0)	2 (0)
1024	1	1	78 (0)	70 (0)	3 (0)	1 (0)
1026	1	2	49 (0)	38 (0)	15 (0)	3 (0)
1027	1	2	52 (0)	34 (0)	3 (0)	0 (0)
1028	1	1	45 (0)	43 (0)	4 (0)	1 (0)

[a] Each number stands for the mean over all experimental sessions with the standard error across sessions in brackets. [b] SUs: Single Units (including ESNs). [c] ESNs: Episode Specific Neurons.

References

Babb, T. L., Carr, E., & Crandall, P. H. (1973). Analysis of extracellular firing patterns of deep temporal lobe structures in man. *Electroencephalogr Clin Neurophysiol, 34*(3), 247-257. https://doi.org/10.1016/0013-4694(73)90252-6

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science, 319*(5870), 1640-1642.

Berens, P. (2009). CircStat: A Matlab Toolbox for Circular Statistics, *Journal of Statistical Software, 31*(10).

Berron, D., Schütze, H., Maass, A., Cardenas-Blanco, A., Kuijf, H. J., Kumaran, D., & Düzel, E. (2016). Strong evidence for pattern separation in human dentate gyrus. *Journal of Neuroscience, 36*(29), 7569-7579.

Berry, S. D., & Thompson, R. F. (1978). Prediction of learning rate from the hippocampal electroencephalogram. *Science, 200*(4347), 1298-1300. https://doi.org/10.1126/science.663612

Bir, S. C., Ambekar, S., Kukreja, S., & Nanda, A. (2015). Julius Caesar Arantius (Giulio Cesare Aranzi, 1530-1589) and the hippocampus of the human brain: history behind the discovery. *Journal of Neurosurgery, 122*(4), 971-975. https://doi.org/10.3171/2014.11.JNS132402

Blatt, M., Wiseman, S., & Domany, E. (1996). Superparamagnetic clustering of data. *Physical review letters, 76*(18), 3251.

Born, J., Rasch, B., & Gais, S. (2006). Sleep to remember. *Neuroscientist, 12*(5), 410-424. https://doi.org/10.1177/1073858406292647

Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological review, 114*(1), 38.

Bush, D., & Burgess, N. (2020). Advantages and detection of phase coding in the absence of rhythmicity. *Hippocampus, 30*(7), 745-762. https://doi.org/10.1002/hipo.23199

Buzsáki, G. (2015). Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus, 25*(10), 1073-1188. https://doi.org/10.1002/hipo.22488

Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents--EEG, ECoG, LFP and spikes. *Nat Rev Neurosci, 13*(6), 407-420. https://doi.org/10.1038/nrn3241

Buzsaki, G., & Tingley, D. (2018). Space and Time: The Hippocampus as a Sequence Generator. *Trends Cogn Sci, 22*(10), 853-869. https://doi.org/10.1016/j.tics.2018.07.006

Cai, D. J., Aharoni, D., Shuman, T., Shobe, J., Biane, J., Song, W., Wei, B., Veshkini, M., La-Vu, M., Lou, J., Flores, S. E., Kim, I., Sano, Y., Zhou, M., Baumgaertel, K., Lavi, A., Kamata, M., Tuszynski, M., Mayford, M., . . . Silva, A. J. (2016). A shared neural ensemble links distinct contextual memories encoded close in time. *Nature, 534*(7605), 115-118. https://doi.org/10.1038/nature17955

Chadwick, M. J., Hassabis, D., Weiskopf, N., & Maguire, E. A. (2010). Decoding individual episodic memory traces in the human hippocampus. *Curr Biol, 20*(6), 544-547. https://doi.org/10.1016/j.cub.2010.01.053

Chaure, F. J., Rey, H. G., & Quian Quiroga, R. (2018). A novel and fully automatic spike-sorting implementation with variable number of features. *Journal of neurophysiology, 120*(4), 1859-1871.

Choi, J. H., Sim, S. E., Kim, J. I., Choi, D. I., Oh, J., Ye, S., Lee, J., Kim, T., Ko, H. G., Lim, C. S., & Kaang, B. K. (2018). Interregional synaptic maps among engram cells underlie memory formation. *Science, 360*(6387), 430-435. https://doi.org/10.1126/science.aas9204

Clouter, A., Shapiro, K. L., & Hanslmayr, S. (2017). Theta Phase Synchronization Is the Glue that Binds Human Associative Memory. *Curr Biol, 27*(20), 3143-3148 e3146. https://doi.org/10.1016/j.cub.2017.09.001

Cohen, M. X. (2017). Multivariate cross-frequency coupling via generalized eigendecomposition. *eLife, 6*. https://doi.org/10.7554/eLife.21792

Cooper, R. A., & Ritchey, M. (2020). Progression from Feature-Specific Brain Activity to Hippocampal Binding during Episodic Encoding. *J Neurosci, 40*(8), 1701-1709. https://doi.org/10.1523/JNEUROSCI.1971-19.2019

Corkin, S. (1984). Lasting Consequences of Bilateral Medial Temporal Lobectomy - Clinical Course and Experimental Findings in Hm. *Seminars in Neurology, 4*(2), 249-259. https://doi.org/10.1055/s-2008-1041556

Corkin, S. (2002). What's new with the amnesic patient H.M.? *Nat Rev Neurosci, 3*(2), 153-160. https://doi.org/10.1038/nrn726

Csicsvari, J., Hirase, H., Czurko, A., Mamiya, A., & Buzsáki, G. (1999). Oscillatory coupling of hippocampal pyramidal cells and interneurons in the behaving Rat. *J Neurosci, 19*(1), 274-287. https://doi.org/10.1523/JNEUROSCI.19-01-00274.1999

De Falco, E., Ison, M. J., Fried, I., & Quian Quiroga, R. (2016). Long-term coding of personal and universal associations underlying the memory web in the human brain. *Nat Commun, 7*, 13408. https://doi.org/10.1038/ncomms13408

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Front Psychol, 5*, 781. https://doi.org/10.3389/fpsyg.2014.00781

Dong, Y., Green, T., Saal, D., Marie, H., Neve, R., Nestler, E. J., & Malenka, R. C. (2006). CREB modulates excitability of nucleus accumbens neurons. *Nat Neurosci, 9*(4), 475-477. https://doi.org/10.1038/nn1661

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A. H., Wallis, J. D., Knight, R. T., Shestyuk, A., & Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nat Neurosci, 23*(12), 1655-1665. https://doi.org/10.1038/s41593-020-00744-x

Durand, S., Heller, G. R., Ramirez, T. K., Luviano, J. A., Williford, A., Sullivan, D. T., Cahoon, A. J., Farrell, C., Groblewski, P. A., Bennett, C., Siegle, J. H., & Olsen, S. R. (2022). Acute head-fixed recordings in awake mice with multiple Neuropixels probes. *Nat Protoc*. https://doi.org/10.1038/s41596-022-00768-6

Dutta, B., Andrei, A., Harris, T. D., Lopez, C. M., O'Callahan, J., Putzeys, J., Raducanu, B. C., Severi, S., Stavisky, S. D., Trautmann, E. M., Welkenhuysen, M., & Shenoy, K. V. (2019). The Neuropixels probe: A CMOS based integrated microsystems platform for neuroscience and brain-computer interfaces. *2019 Ieee International Electron Devices Meeting (Iedm)*.

Engel, J., Jr. (1996). Surgery for seizures. *N Engl J Med, 334*(10), 647-652. https://doi.org/10.1056/NEJM199603073341008

Frankland, P. W., & Josselyn, S. A. (2015). Memory allocation. *Neuropsychopharmacology, 40*(1), 243. https://doi.org/10.1038/npp.2014.234

Fried, I., Wilson, C. L., Maidment, N. T., Engel, J., Behnke, E., Fields, T. A., MacDonald, K. A., Morrow, J. W., & Ackerson, L. (1999). Cerebral microdialysis combined with single-neuron and electroencephalographic recording in neurosurgical patients - Technical note. *Journal of Neurosurgery, 91*(4), 697-705. https://doi.org/DOI 10.3171/jns.1999.91.4.0697

Galvani, Luigi. (1791). *De viribus electricitatis in motu musculari commentarius.* 7, 363-418.

Gao, J., Hu, J., & Tung, W. W. (2011). Facilitating joint chaos and fractal analysis of biosignals through nonlinear adaptive filtering. *PloS one, 6*(9), e24331. https://doi.org/10.1371/journal.pone.0024331

Gao, R., Peterson, E. J., & Voytek, B. (2017). Inferring synaptic excitation/inhibition balance from field potentials. *Neuroimage, 158*, 70-78. https://doi.org/10.1016/j.neuroimage.2017.06.078

Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., & Fried, I. (2008). Internally generated reactivation of single neurons in human hippocampus during free recall. *Science, 322*(5898), 96-101.

Goyal, A., Miller, J., Qasim, S. E., Watrous, A. J., Zhang, H., Stein, J. M., Inman, C. S., Gross, R. E., Willie, J. T., Lega, B., Lin, J. J., Sharan, A., Wu, C., Sperling, M. R., Sheth, S. A., McKhann, G. M., Smith, E. H., Schevon, C., & Jacobs, J. (2020). Functionally distinct high and low theta oscillations in the human hippocampus. *Nat Commun, 11*(1), 2469. https://doi.org/10.1038/s41467-020-15670-6

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica, 37*(3), 424-438. https://doi.org/Doi 10.2307/1912791

Griffiths, B. J., Martin-Buro, M. C., Staresina, B. P., & Hanslmayr, S. (2021). Disentangling neocortical alpha/beta and hippocampal theta/gamma oscillations in human episodic memory formation. *Neuroimage, 242*, 118454. https://doi.org/10.1016/j.neuroimage.2021.118454

Hafting, T., Fyhn, M., Molden, S., Moser, M. B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature, 436*(7052), 801-806. https://doi.org/10.1038/nature03721

Hampson, R. E., Song, D., Robinson, B. S., Fetterhoff, D., Dakos, A. S., Roeder, B. M., She, X., Wicks, R. T., Witcher, M. R., Couture, D. E., Laxton, A. W., Munger-Clary, H., Popli, G., Sollman, M. J., Whitlow, C. T., Marmarelis, V. Z., Berger, T. W., & Deadwyler, S. A. (2018). Developing a hippocampal neural prosthetic to facilitate human memory encoding and recall. *J Neural Eng, 15*(3), 036014. https://doi.org/10.1088/1741-2552/aaaed7

Han, J. H., Kushner, S. A., Yiu, A. P., Hsiang, H. L., Buch, T., Waisman, A., Bontempi, B., Neve, R. L., Frankland, P. W., & Josselyn, S. A. (2009). Selective erasure of a fear memory. *Science, 323*(5920), 1492-1496. https://doi.org/10.1126/science.1164139

Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C. S., & Bauml, K. H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. *Neuroimage, 37*(4), 1465-1473. https://doi.org/10.1016/j.neuroimage.2007.07.011

Hanslmayr, S., Staresina, B. P., & Bowman, H. (2016). Oscillations and Episodic Memory: Addressing the Synchronization/Desynchronization Conundrum. *Trends Neurosci, 39*(1), 16-25. https://doi.org/10.1016/j.tins.2015.11.004

Hardt, O., Nader, K., & Nadel, L. (2013). Decay happens: the role of active forgetting in memory. *Trends Cogn Sci, 17*(3), 111-120. https://doi.org/10.1016/j.tics.2013.01.001

Hasselmo, M. E., Bodelon, C., & Wyble, B. P. (2002). A proposed function for hippocampal theta rhythm: separate phases of encoding and retrieval enhance reversal of prior learning. *Neural Comput, 14*(4), 793-817. https://doi.org/10.1162/089976602317318965

Hebb, D. O. (1949). *The Organization of behavior: A Neuropsychological Theory*. Wiley and Sonds.

Herweg, N. A., Solomon, E. A., & Kahana, M. J. (2020). Theta Oscillations in Human Memory. *Trends Cogn Sci, 24*(3), 208-227. https://doi.org/10.1016/j.tics.2019.12.006

Hirsh, R. (1974). The hippocampus and contextual retrieval of information from memory: a theory. *Behav Biol, 12*(4), 421-444. https://doi.org/10.1016/s0091-6773(74)92231-7

Huxter, J., Burgess, N., & O'Keefe, J. (2003). Independent rate and temporal coding in hippocampal pyramidal cells. *Nature, 425*(6960), 828-832. https://doi.org/10.1038/nature02058

Ison, M. J., Quian Quiroga, R., & Fried, I. (2015). Rapid encoding of new memories by individual neurons in the human brain. *Neuron, 87*(1), 220-230.

Jacobs, J., Kahana, M. J., Ekstrom, A. D., & Fried, I. (2007). Brain oscillations control timing of single-neuron activity in humans. *J Neurosci, 27*(14), 3839-3844. https://doi.org/10.1523/JNEUROSCI.4636-06.2007

Jadhav, S. P., Kemere, C., German, P. W., & Frank, L. M. (2012). Awake hippocampal sharp-wave ripples support spatial memory. *Science, 336*(6087), 1454-1458. https://doi.org/10.1126/science.1217230

Jensen, O., Kaiser, J., & Lachaux, J. P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci, 30*(7), 317-324. https://doi.org/10.1016/j.tins.2007.05.001

Josselyn, S. A. (2010). Continuing the search for the engram: examining the mechanism of fear memories. *J Psychiatry Neurosci, 35*(4), 221-228. https://doi.org/10.1503/jpn.100015

Josselyn, S. A., Kohler, S., & Frankland, P. W. (2015). Finding the engram. *Nat Rev Neurosci, 16*(9), 521-534. https://doi.org/10.1038/nrn4000

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., Lee, A. K., Anastassiou, C. A., Andrei, A., Aydin, C., Barbic, M., Blanche, T. J., Bonin, V., Couto, J., Dutta, B., Gratiy, S. L., Gutnisky, D. A., Hausser, M., Karsh, B., . . . Harris, T. D. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature, 551*(7679), 232-236. https://doi.org/10.1038/nature24636

Kerrén, C., Linde-Domingo, J., Hanslmayr, S., & Wimber, M. (2018). An Optimal Oscillatory Phase for Pattern Reactivation during Memory Retrieval. *Curr Biol, 28*(21), 3383-3392 e3386. https://doi.org/10.1016/j.cub.2018.08.065

Kerrén, C., van Bree, S., Griffiths, B. J., & Wimber, M. (2022). Phase separation of competing memories along the human hippocampal theta rhythm. e*Life, 11*. https://doi.org/10.7554/eLife.80633

Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res Rev, 53*(1), 63-88. https://doi.org/10.1016/j.brainresrev.2006.06.003

Kolibius, L. D., Born, J., & Feld, G. B. (2020). Vast Amounts of Encoded Items Nullify but Do Not Reverse the Effect of Sleep on Declarative Memory. *Front Psychol, 11*, 607070. https://doi.org/10.3389/fpsyg.2020.607070

Kota, S., Rugg, M. D., & Lega, B. C. (2020). Hippocampal Theta Oscillations Support Successful Associative Memory Formation. *J Neurosci, 40*(49), 9507-9518. https://doi.org/10.1523/JNEUROSCI.0767-20.2020

Krotov, D., & Hopfield, J. (2020). Large associative memory problem in neurobiology and machine learning. *arXiv preprint*.

Kucewicz, M. T., Cimbalnik, J., Matsumoto, J. Y., Brinkmann, B. H., Bower, M. R., Vasoli, V., Sulc, V., Meyer, F., Marsh, W. R., Stead, S. M., & Worrell, G. A. (2014). High frequency oscillations are associated with cognitive processing in human recognition memory. *Brain, 137*(Pt 8), 2231-2244. https://doi.org/10.1093/brain/awu149

Kunz, L., Brandt, A., Reinacher, P. C., Staresina, B. P., Reifenstein, E. T., Weidemann, C. T., Herweg, N. A., Patel, A., Tsitsiklis, M., Kempter, R., Kahana, M. J., Schulze-Bonhage, A., & Jacobs, J. (2021). A neural code for egocentric spatial maps in the human medial temporal lobe. *Neuron, 109*(17), 2781-2796 e2710. https://doi.org/10.1016/j.neuron.2021.06.019

Kwan, P., Schachter, S. C., & Brodie, M. J. (2011). Drug-resistant epilepsy. *N Engl J Med, 365*(10), 919-926. https://doi.org/10.1056/NEJMra1004418

Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Curr Biol, 22*(11), 1000-1004. https://doi.org/10.1016/j.cub.2012.03.054

Lega, B. C., Jacobs, J., & Kahana, M. (2012). Human hippocampal theta oscillations and the formation of episodic memories. *Hippocampus, 22*(4), 748-761. https://doi.org/10.1002/hipo.20937

Leszczyński, M., Barczak, A., Kajikawa, Y., Ulbert, I., Falchier, A. Y., Tal, I., Haegens, S., Melloni, L., Knight, R. T., & Schroeder, C. E. (2020). Dissociation of broadband high-frequency activity and neuronal firing in the neocortex. *Sci Adv, 6*(33), eabb0977. https://doi.org/10.1126/sciadv.abb0977

Lifanov, J., Griffiths, B. J., Linde-Domingo, J., Ferreira, C. S., Wilson, M., Mayhew, S. D., Charest, I., & Wimber, M. (2022). Reconstructing Spatio-Temporal Trajectories of Visual Object Memories in the Human Brain. *bioRxiv preprint*. https://doi.org/10.1101/2022.12.15.520591

Linde-Domingo, J., Treder, M. S., Kerren, C., & Wimber, M. (2019). Evidence that neural information flow is reversed between object perception and object reconstruction from memory. *Nat Commun, 10*(1), 179. https://doi.org/10.1038/s41467-018-08080-2

Linden, H., Pettersen, K. H., & Einevoll, G. T. (2010). Intrinsic dendritic filtering gives low-pass power spectra of local field potentials. *J Comput Neurosci, 29*(3), 423-444. https://doi.org/10.1007/s10827-010-0245-4

Lisman, J., Buzsáki, G., Eichenbaum, H., Nadel, L., Ranganath, C., & Redish, A. D. (2017). Viewpoints: how the hippocampus contributes to memory, navigation and cognition. *Nature neuroscience, 20*(11), 1434-1447.

Liu, X., Ramirez, S., Pang, P. T., Puryear, C. B., Govindarajan, A., Deisseroth, K., & Tonegawa, S. (2012). Optogenetic stimulation of a hippocampal engram activates fear memory recall. *Nature, 484*(7394), 381-385. https://doi.org/10.1038/nature11028

Mack, M. L., & Preston, A. R. (2016). Decisions about the past are guided by reinstatement of specific memories in the hippocampus and perirhinal cortex. *Neuroimage, 127*, 144-157.

Manning, J. R., Jacobs, J., Fried, I., & Kahana, M. J. (2009). Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. *J Neurosci, 29*(43), 13613-13620. https://doi.org/10.1523/JNEUROSCI.2041-09.2009

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods, 164*(1), 177-190.

Marr, D. (1971). Simple memory: a theory for archicortex. *Philos Trans R Soc Lond B Biol Sci, 262*(841), 23-81. https://doi.org/10.1098/rstb.1971.0078

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev, 102*(3), 419-457. https://doi.org/10.1037/0033-295X.102.3.419

Miller, K. J., Sorensen, L. B., Ojemann, J. G., & den Nijs, M. (2009). Power-law scaling in the brain surface electric potential. *PLoS Comput Biol, 5*(12), e1000609. https://doi.org/10.1371/journal.pcbi.1000609

Mormann, F., Dubois, J., Kornblith, S., Milosavljevic, M., Cerf, M., Ison, M., Tsuchiya, N., Kraskov, A., Quiroga, R. Q., & Adolphs, R. (2011). A category-specific response to animals in the right human amygdala. *Nature neuroscience, 14*(10), 1247-1249.

Mormann, F., Kornblith, S., Quian Quiroga, R., Kraskov, A., Cerf, M., Fried, I., & Koch, C. (2008). Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *Journal of Neuroscience, 28*(36), 8865-8872.

Murphy, S. L., Kochanek, K. D., Xu, J. Q., & Arias, E. (2021). *Mortality in the United States, 2020.* [NCHS Data Brief.]. https://www.cdc.gov/nchs/data/databriefs/db427.pdf

Naber, P. A., Lopes da Silva, F. H., & Witter, M. P. (2001). Reciprocal connections between the entorhinal cortex and hippocampal fields CA1 and the subiculum are in register with the projections from CA1 to the subiculum. *Hippocampus, 11*(2), 99-104. https://doi.org/10.1002/hipo.1028

Nadel, L., & Moscovitch, M. (1997). Memory consolidation, retrograde amnesia and the hippocampal complex. *Curr Opin Neurobiol, 7*(2), 217-227. https://doi.org/10.1016/s0959-4388(97)80010-4

Nader, K., & Hardt, O. (2009). A single standard for memory: the case for reconsolidation. *Nat Rev Neurosci, 10*(3), 224-234. https://doi.org/10.1038/nrn2590

Ngo, H. V., Fell, J., & Staresina, B. (2020). Sleep spindles mediate hippocampal-neocortical coupling during long-duration ripples. *eLife, 9*. https://doi.org/10.7554/eLife.57011

Niediek, J., Boström, J., Elger, C. E., & Mormann, F. (2016). Reliable analysis of single-unit recordings from the human brain under noisy conditions: tracking neurons over hours. *PloS one, 11*(12), e0166598.

Nir, Y., Fisch, L., Mukamel, R., Gelbard-Sagiv, H., Arieli, A., Fried, I., & Malach, R. (2007). Coupling between neuronal firing rate, gamma LFP, and BOLD fMRI is related to interneuronal correlations. *Curr Biol, 17*(15), 1275-1285. https://doi.org/10.1016/j.cub.2007.06.066

Nolte, G., Ziehe, A., Nikulin, V. V., Schlogl, A., Kramer, N., Brismar, T., & Muller, K. R. (2008). Robustly estimating the flow direction of information in complex physical systems. *Physical review letters, 100*(23). https://doi.org/10.1103/PhysRevLett.100.234101

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res, 34*(1), 171-175. https://doi.org/10.1016/0006-8993(71)90358-1

O'Keefe, J., & Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus, 3*(3), 317-330. https://doi.org/10.1002/hipo.450030307

O'Reilly, R. C., & Rudy, J. W. (2001). Conjunctive representations in learning and memory: principles of cortical and hippocampal function. *Psychol Rev, 108*(2), 311-345. https://doi.org/10.1037/0033-295x.108.2.311

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci*. https://doi.org/10.1155/2011/156869

Pacheco Estefan, D., Sanchez-Fibla, M., Duff, A., Principe, A., Rocamora, R., Zhang, H., Axmacher, N., & Verschure, P. (2019). Coordinated representational reinstatement in the human hippocampus and lateral temporal cortex during episodic memory retrieval. *Nat Commun, 10*(1), 2255. https://doi.org/10.1038/s41467-019-09569-0

Pachitariu, M., Sridhar, S., & Stringer, C. (2023). Solving the spike sorting problem with Kilosort. *bioRxiv preprint*. https://doi.org/10.1101/2023.01.07.523036

Parish, G., Hanslmayr, S., & Bowman, H. (2018). The Sync/deSync Model: How a Synchronized Hippocampus and a Desynchronized Neocortex Code Memories. *J Neurosci, 38*(14), 3428-3440. https://doi.org/10.1523/JNEUROSCI.2561-17.2018

Parish, G., Michelmann, S., Hanslmayr, S., & Bowman, H. (2021). The Sync-Fire/deSync model: Modelling the reactivation of dynamic memories from cortical alpha oscillations. *Neuropsychologia, 158*, 107867. https://doi.org/10.1016/j.neuropsychologia.2021.107867

Park, S., Kramer, E. E., Mercaldo, V., Rashid, A. J., Insel, N., Frankland, P. W., & Josselyn, S. A. (2016). Neuronal Allocation to a Hippocampal Engram. *Neuropsychopharmacology, 41*(13), 2987-2993. https://doi.org/10.1038/npp.2016.73

Parvizi, J., & Kastner, S. (2018). Promises and limitations of human intracranial electroencephalography. *Nat Neurosci, 21*(4), 474-483. https://doi.org/10.1038/s41593-018-0108-2

Paulk, A. C., Kfir, Y., Khanna, A. R., Mustroph, M. L., Trautmann, E. M., Soper, D. J., Stavisky, S. D., Welkenhuysen, M., Dutta, B., Shenoy, K. V., Hochberg, L. R., Richardson, R. M., Williams, Z. M., & Cash, S. S. (2022). Large-scale neural recordings with single neuron resolution using Neuropixels probes in human cortex. *Nat Neurosci, 25*(2), 252-263. https://doi.org/10.1038/s41593-021-00997-0

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., & Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images*. Elsevier.

Piccolino, M. (1998). Animal electricity and the birth of electrophysiology: the legacy of Luigi Galvani. *Brain Res Bull, 46*(5), 381-407. https://doi.org/10.1016/s0361-9230(98)00026-4

Prestigio, C., Ferrante, D., Valente, P., Casagrande, S., Albanesi, E., Yanagawa, Y., Benfenati, F., & Baldelli, P. (2019). Spike-Related Electrophysiological Identification of Cultured Hippocampal Excitatory and Inhibitory Neurons. *Mol Neurobiol, 56*(9), 6276-6292. https://doi.org/10.1007/s12035-019-1506-5

Quian Quiroga, R. (2012). Concept cells: the building blocks of declarative memory functions. *Nat Rev Neurosci, 13*(8), 587-597. https://doi.org/10.1038/nrn3251

Quian Quiroga, R. (2019). Plugging in to Human Memory: Advantages, Challenges, and Insights from Human Single-Neuron Recordings. *Cell, 179*(5), 1015-1032. https://doi.org/10.1016/j.cell.2019.10.016

Quian Quiroga, R. (2020). No pattern separation in the human hippocampus. *Trends in cognitive sciences, 24*(12), 994-1007.

Quian Quiroga, R., Kraskov, A., Mormann, F., Fried, I., & Koch, C. (2014). Single-cell responses to face adaptation in the human medial temporal lobe. *Neuron, 84*(2), 363-369. https://doi.org/10.1016/j.neuron.2014.09.006

Quian Quiroga, R., Mukamel, R., Isham, E. A., Malach, R., & Fried, I. (2008). Human single-neuron responses at the threshold of conscious recognition. Proc Natl Acad Sci U S A, 105(9), 3599-3604. https://doi.org/10.1073/pnas.0707043105

Quian Quiroga, R., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature, 435*(7045), 1102-1107.

Ramirez, S., Liu, X., Lin, P. A., Suh, J., Pignatelli, M., Redondo, R. L., Ryan, T. J., & Tonegawa, S. (2013). Creating a false memory in the hippocampus. *Science, 341*(6144), 387-391. https://doi.org/10.1126/science.1239073

Ramón y Cajal, S. (1888). Estructura de los centros nerviosos de las aves. *Revista trimestral de histología normal y patológica, 1*, 1-10.

Ramón y Cajal, S. (1894). The Croonian Lecture. La fine structure des centres nerveux. *Proceedings of the Royal Society B: Biological Sciences, 55*, 444-468.

Rao-Ruiz, P., Yu, J., Kushner, S. A., & Josselyn, S. A. (2019). Neuronal competition: microcircuit mechanisms define the sparsity of the engram. *Curr Opin Neurobiol, 54*, 163-170. https://doi.org/10.1016/j.conb.2018.10.013

Rashid, A. J., Yan, C., Mercaldo, V., Hsiang, H. L., Park, S., Cole, C. J., De Cristofaro, A., Yu, J., Ramakrishnan, C., Lee, S. Y., Deisseroth, K., Frankland, P. W., & Josselyn, S. A. (2016). Competition between engrams influences fear memory formation and recall. *Science, 353*(6297), 383-387. https://doi.org/10.1126/science.aaf0594

Ray, S., Crone, N. E., Niebur, E., Franaszczuk, P. J., & Hsiao, S. S. (2008). Neural correlates of high-gamma oscillations (60-200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *J Neurosci, 28*(45), 11526-11536. https://doi.org/10.1523/JNEUROSCI.2848-08.2008

Ray, S., & Maunsell, J. H. (2011). Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol, 9*(4), e1000610. https://doi.org/10.1371/journal.pbio.1000610

Reddy, L., Self, M. W., Zoefel, B., Poncet, M., Possel, J. K., Peters, J. C., Baayen, J. C., Idema, S., VanRullen, R., & Roelfsema, P. R. (2021). Theta-phase dependent neuronal coding during sequence learning in human single neurons. *Nat Commun, 12*(1), 4839. https://doi.org/10.1038/s41467-021-25150-0

Reddy, L., Zoefel, B., Possel, J. K., Peters, J., Dijksterhuis, D. E., Poncet, M., van Straaten, E. C., Baayen, J. C., Idema, S., & Self, M. W. (2021). Human hippocampal neurons track moments in a sequence of events. *Journal of Neuroscience, 41*(31), 6714-6725.

Redish, A. D., Battaglia, F. P., Chawla, M. K., Ekstrom, A. D., Gerrard, J. L., Lipa, P., Rosenzweig, E. S., Worley, P. F., Guzowski, J. F., McNaughton, B. L., & Barnes, C. A. (2001). Independence of firing correlates of anatomically proximate hippocampal pyramidal cells. *J Neurosci, 21*(5), RC134. https://doi.org/10.1523/JNEUROSCI.21-05-j0004.2001

Rødland, E. A. (2006). Simes' procedure is ‘valid on average’. Biometrika, 93(3), 742-746.

Roux, F., Parish, G., Chelvarajah, R., Rollings, D. T., Sawlani, V., Hamer, H., Gollwitzer, S., Kreiselmeyer, G., Ter Wal, M. J., Kolibius, L., Staresina, B. P., Wimber, M., Self, M. W., & Hanslmayr, S. (2022). Oscillations support short latency co-firing of neurons during human episodic memory formation. *eLife, 11*. https://doi.org/10.7554/eLife.78109

Roux, L., Hu, B., Eichler, R., Stark, E., & Buzsáki, G. (2017). Sharp wave ripples during learning stabilize the hippocampal spatial map. *Nat Neurosci, 20*(6), 845-853. https://doi.org/10.1038/nn.4543

Roy, D. S., Park, Y.-G., Ogawa, S. K., Cho, J. H., Choi, H., Kamensky, L., Martin, J., Chung, K., & Tonegawa, S. (2019). Brain-wide mapping of contextual fear memory engram ensembles supports the dispersed engram complex hypothesis. *bioRxiv preprint*, 668483. https://doi.org/10.1101/668483

Rutishauser, U., Mamelak, A. N., & Schuman, E. M. (2006). Single-trial learning of novel stimuli by individual neurons of the human hippocampus-amygdala complex. *Neuron, 49*(6), 805-813. https://doi.org/10.1016/j.neuron.2006.02.015

Rutishauser, U., Ross, I. B., Mamelak, A. N., & Schuman, E. M. (2010). Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature, 464*(7290), 903-907. https://doi.org/10.1038/nature08860

Rutishauser, U., Schuman, E. M., & Mamelak, A. N. (2008). Activity of human hippocampal and amygdala neurons during retrieval of declarative memories. *Proceedings of the National Academy of Sciences, 105*(1), 329-334.

Rutishauser, U., Ye, S., Koroma, M., Tudusciuc, O., Ross, I. B., Chung, J. M., & Mamelak, A. N. (2015). Representation of retrieval confidence by single neurons in the human medial temporal lobe. *Nature neuroscience, 18*(7), 1041-1050.

Samaha, J., & Cohen, M. X. (2022). Power spectrum slope confounds estimation of instantaneous oscillatory frequency. *Neuroimage, 250*, 118929. https://doi.org/10.1016/j.neuroimage.2022.118929

Samuelson, K. W. (2011). Post-traumatic stress disorder and declarative memory functioning: a review. *Dialogues Clin Neurosci, 13*(3), 346-351. https://doi.org/10.31887/DCNS.2011.13.2/ksamuelson

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Philos Trans R Soc Lond B Biol Sci, 372*(1711). https://doi.org/10.1098/rstb.2016.0049

Schreiner, T., Petzka, M., Staudigl, T., & Staresina, B. P. (2021). Endogenous memory reactivation during sleep in humans is clocked by slow oscillation-spindle complexes. *Nat Commun, 12*(1), 3112. https://doi.org/10.1038/s41467-021-23520-2

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of neurology, neurosurgery, and psychiatry, 20*(1), 11.

Semon, R. W. (1904). *Die Mneme : als erhaltendes Prinzip im Wechsel des organischen Geschehens.* Wilhelm Engelmann.

Shatz, C. J. (1992). The developing brain. *Sci Am, 267*(3), 60-67. https://doi.org/10.1038/scientificamerican0992-60

Silva, A. J., Zhou, Y., Rogerson, T., Shobe, J., & Balaji, J. (2009). Molecular and cellular approaches to memory allocation in neural circuits. *Science, 326*(5951), 391-395. https://doi.org/10.1126/science.1174519

Squire, L. R. (1987). *Memory and brain.* Oxford University Press.

Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological review, 99*(2), 195.

Squire, L. R. (2009). The legacy of patient H.M. for neuroscience. *Neuron, 61*(1), 6-9. https://doi.org/10.1016/j.neuron.2008.12.023

Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: a neurobiological perspective. *Curr Opin Neurobiol, 5*(2), 169-177. https://doi.org/10.1016/0959-4388(95)80023-9

Swanson, L. W., & Mogenson, G. J. (1981). Neural mechanisms for the functional coupling of autonomic, endocrine and somatomotor responses in adaptive behavior. *Brain Res, 228*(1), 1-34. https://doi.org/10.1016/0165-0173(81)90010-2

Tamamaki, N., & Nojyo, Y. (1995). Preservation of topography in the connections between the subiculum, field CA1, and the entorhinal cortex in rats. *J Comp Neurol, 353*(3), 379-390. https://doi.org/10.1002/cne.903530306

Tankus, A., Yeshurun, Y., & Fried, I. (2009). An automatic measure for classifying clusters of suspected spikes into single cells versus multiunits. *Journal of neural engineering, 6*(5), 056001.

Taube, J. S., Muller, R. U., & Ranck, J. B., Jr. (1990). Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *J Neurosci, 10*(2), 420-435. https://doi.org/10.1523/JNEUROSCI.10-02-00420.1990

ter Wal, M., Linde-Domingo, J., Lifanov, J., Roux, F., Kolibius, L. D., Gollwitzer, S., Lang, J., Hamer, H., Rollings, D., Sawlani, V., Chelvarajah, R., Staresina, B., Hanslmayr, S., & Wimber, M. (2021). Theta rhythmicity governs human behavior and hippocampal signals during memory-dependent tasks. *Nat Commun, 12*(1), 7048. https://doi.org/10.1038/s41467-021-27323-3

Teyler, T. J., & DiScenna, P. (1986). The hippocampal memory indexing theory. *Behavioral neuroscience, 100*(2), 147.

Teyler, T. J., & Rudy, J. W. (2007). The hippocampal indexing theory and episodic memory: updating the index. *Hippocampus, 17*(12), 1158-1169.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual review of psychology, 53*(1), 1-25.
Tulving, E. D., W (Ed.). (1972). *Episodic and semantic memory.* Academic Press.
Umbach, G., Kantak, P., Jacobs, J., Kahana, M., Pfeiffer, B. E., Sperling, M., & Lega, B. (2020). Time cells in the human hippocampus and entorhinal cortex support episodic memory. *Proceedings of the National Academy of Sciences, 117*(45), 28463-28474.
Underwood, B. J. (1957). Interference and forgetting. *Psychol Rev, 64*(1), 49-60. https://doi.org/10.1037/h0044616
Vaz, A. P., Inati, S. K., Brunel, N., & Zaghloul, K. A. (2019). Coupled ripple oscillations between the medial temporal lobe and neocortex retrieve human memory. *Science, 363*(6430), 975-978. https://doi.org/10.1126/science.aau8956
Vetere, G., Tran, L. M., Moberg, S., Steadman, P. E., Restivo, L., Morrison, F. G., Ressler, K. J., Josselyn, S. A., & Frankland, P. W. (2019). Memory formation in the absence of experience. *Nat Neurosci, 22*(6), 933-940. https://doi.org/10.1038/s41593-019-0389-0
Ward, A. A., & Thomas, L. B. (1955). The electrical activity of single units in the cerebral cortex of man. *Electroencephalogr Clin Neurophysiol, 7*(1), 135-136. https://doi.org/10.1016/0013-4694(55)90067-5
Whittingstall, K., & Logothetis, N. K. (2009). Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex. *Neuron, 64*(2), 281-289. https://doi.org/10.1016/j.neuron.2009.08.016
Whittington, J. C. R., Warren, J., & Behrens, T. E. J. (2022). Relating transformers to models and neural representations of the hippocampal formation. *arXiv preprint*.
Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science, 322*(5901), 606-607. https://doi.org/10.1126/science.1162548
Winson, J. (1978). Loss of hippocampal theta rhythm results in spatial memory deficit in the rat. *Science, 201*(4351), 160-163. https://doi.org/10.1126/science.663646
Wixted, J. T., Goldinger, S. D., Squire, L. R., Kuhn, J. R., Papesh, M. H., Smith, K. A., Treiman, D. M., & Steinmetz, P. N. (2018). Coding of episodic memory in the human hippocampus. *Proceedings of the National Academy of Sciences, 115*(5), 1093-1098.
Wixted, J. T., Squire, L. R., Jang, Y., Papesh, M. H., Goldinger, S. D., Kuhn, J. R., Smith, K. A., Treiman, D. M., & Steinmetz, P. N. (2014). Sparse and distributed coding of episodic memory in neurons of the human hippocampus. *Proceedings of the National Academy of Sciences, 111*(26), 9621-9626.
World Health Organization. (2019). *ICD-11: International classification of diseases* (11th revision). https://icd.who.int/
World Health Organization. (2022). *Dementia.* Retrieved 31.01.2023 from https://www.who.int/news-room/fact-sheets/detail/dementia
Yokose, J., Okubo-Suzuki, R., Nomoto, M., Ohkawa, N., Nishizono, H., Suzuki, A., Matsuo, M., Tsujimura, S., Takahashi, Y., Nagase, M., Watabe, A. M., Sasahara, M., Kato, F., & Inokuchi, K. (2017). Overlapping memory trace indispensable for linking, but not recalling, individual memories. *Science, 355*(6323), 398-403. https://doi.org/10.1126/science.aal2690
Zheng, J., Schjetnan, A. G. P., Yebra, M., Gomes, B. A., Mosher, C. P., Kalia, S. K., Valiante, T. A., Mamelak, A. N., Kreiman, G., & Rutishauser, U. (2022). Neurons detect cognitive boundaries to structure episodic memories in humans. *Nat Neurosci, 25*(3), 358-368. https://doi.org/10.1038/s41593-022-01020-w