



Nikolai Axmacher

Abteilung für Neuropsychologie, Institut für Kognitive Neurowissenschaft, Fakultät für Psychologie, Ruhr-Universität Bochum, Bochum, Germany

In search of the human engram

What is an engram?

Memory is not only an important cognitive function; it ensures the continuity of our self-image and therefore our identity. We “are” our memories — this is why we feel fundamentally threatened by diseases like Alzheimer’s dementia in which we lose our memory functions. Numerous psychological studies have investigated memory processes already, and much is known about the neuroscientific basis of learning and memory. However, most studies have focused on investigating the general processes of memory rather than on identifying individual memory traces. For example, studies have explored which conditions facilitate memory for faces, which brain systems are relevant for learning to play an instrument, and how facts are stored in comparison to episodes. Only a few studies have investigated how *this* specific situation, or the ability to play *that* song, are represented in the brain and converted into enduring memory traces; how a memory of a friend reactivates a very specific image of that person; or why my contribution to a specific theatre play at school may appear increasingly brilliant as I grow older. These questions refer to a central concept that for a long time appeared to be the “holy grail” of memory research, but for which observation in humans appeared infeasible: the engram. An engram is the unique trace of an experience in the brain, which differs from the trace of all other experiences. In his classic 1950 work, “In search of the engram” [6], Karl Spencer Lashley summarized a large number of lesion studies that aimed at localizing engrams. Despite extensive lesions, surprisingly few memory deficits were observed, leading

Lashley to come to the pessimistic conclusion: “It is not possible to demonstrate the isolated localization of a memory trace anywhere within the nervous system.”

Only recently has the dream to directly observe an engram re-emerged as a reality — and even in humans, this now appears to be within reach. On several levels of brain organization, from the firing of individual nerve cells via the specific activity of small- and medium-scale networks, to activity patterns within widely distributed brain areas, it has been possible to identify elements of specific memory traces in humans. Furthermore, preliminary evidence suggests that content-specific representations — engrams — are particularly impaired in Alzheimer’s disease and may be restored via novel therapeutic methods.

Engrams: from neurons to networks

Despite their complexity, nerve cells are considered the elementary unit of information processing in the brain. In humans it is usually not possible to record the activity of individual neurons. However, there is one exception: recordings in epilepsy patients who have been implanted with microwires during presurgical planning. It has been shown that the activity of individual cells in the hippocampus increases at specific spatial locations when a virtual reality task is used to enable these patients to pseudonavigate themselves to these locations [4]. This activity is thought to correspond to “place cells” in rodents. Other hippocampal cells have been shown to react specifically to the presentation of an image of a particular person ([7]; ■ Fig. 1). These representations show a remarkable de-

gree of perceptual invariance, and these cells have therefore been labeled “concept cells”. Due to their content specificity, both human *place cells* and *concept cells* could constitute the cellular basis of engrams. However, even in relatively sparse hippocampal representations, individual contents are not represented by only one neuron but by hundreds of thousands, or even millions, of cells; and each single cell in the hippocampus seems to be part of not only one, but several of these networks. While these investigations are allowing for fascinating new insights into the neural foundation of engrams on a cellular level, many open questions remain. For example, we still know very little about the mechanisms that determine which concept is represented by which cell; the processes governing the recruitment of neurons into the relevant networks during memory formation; or the changes of representations through synaptic plasticity.

As content-specific cells are always part of larger networks, it appears feasible to also detect engrams on the level of networks. The activity of relatively small neural networks within the human hippocampus (or in other areas such as the neocortex), on the level of local field potentials, can also be investigated via microelectrodes. More extended networks can again be explored in epilepsy patients during the presurgical planning stage, using intracranial EEG recordings with larger electrodes (with a diameter of 1–1.5 mm). The activity of these networks is characterized by a highly specific pattern of rhythmic and arrhythmic activity across different frequencies (■ Fig. 2). On a very general level, the activity of a network reflects the excitability of its constituent neurons. The exact ac-

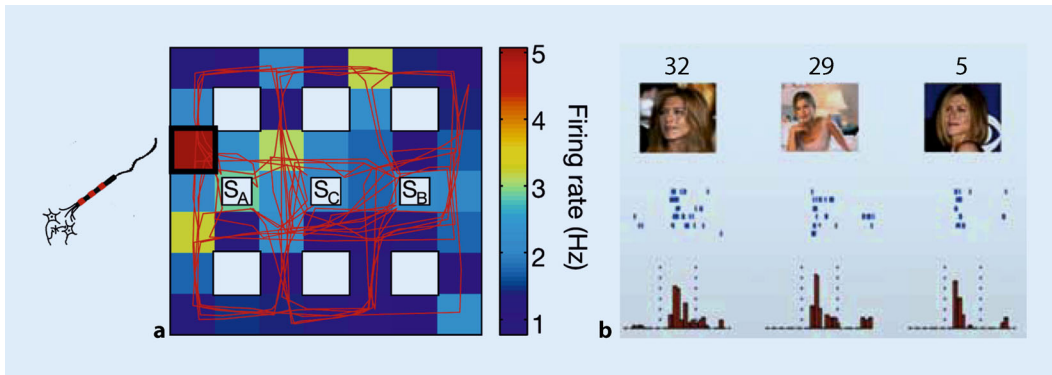


Fig. 1 ▲ Stimulus-specific representations in individual cells. Single-unit recordings show increased activity of individual cells in the human hippocampus at a specific location during spatial navigation (a) or during presentation of a specific person (b). a Color-coded rate of action potentials depending on spatial position. b *Top*, presented image; *middle*: action potentials during repeated presentation of images; *bottom*: histogram of action potentials. Figures reproduced with permission from [4] (a) and [7] (b)

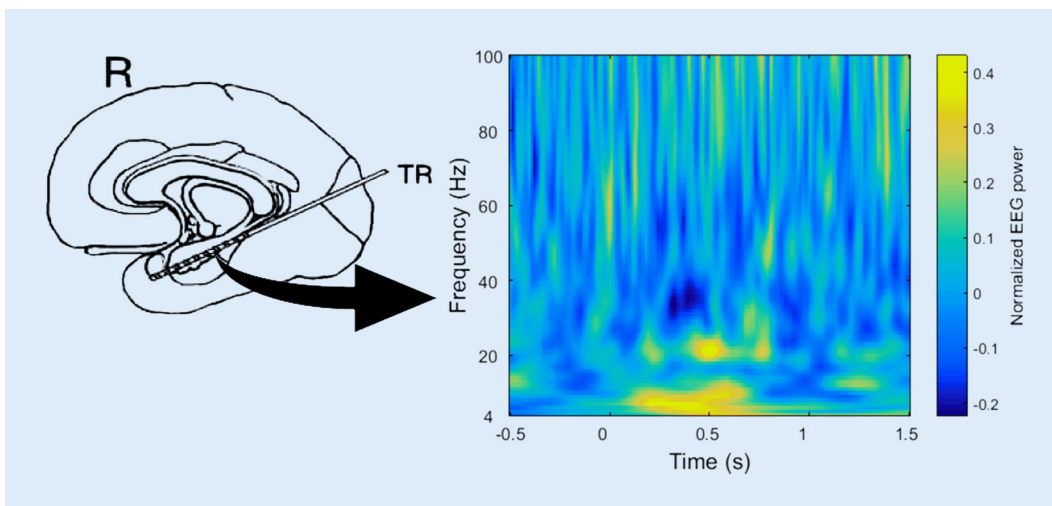


Fig. 2 ◀ Time–frequency pattern of activity in neural networks. Intracranial EEG recordings in epilepsy patients (*left*, schematic figure of a hippocampal depth electrode) allow measuring the time–frequency pattern of neural networks even in deep brain areas such as the hippocampus (*right*). Figure by Hui Zhang

tivity pattern depends on a large number of anatomical and physiological parameters, such as the number of contributing cells, their activity state, and their connectivity. This pattern can be dynamically adjusted in order to perform the cognitive operations and fulfil the behavioral demands that are required in a given situation.

Recent studies have shown that the functional state of a network, reflected via its intracranial EEG time–frequency pattern, does not only play an important general role for memory formation, but differs depending on the processed content and could therefore constitute a basis of engrams on the network level. For example, it has been possible via intracranial EEG recordings to identify the representations of individual letters [9] or spatial locations ([10]; ■ Fig. 3). These

studies were based on the application of multivariate pattern classification algorithms (MVPA) that had been initially developed in the computer sciences and for artificial intelligence research. These methods have allowed for the “decoding” of specific neural contents based on network-level activity. Such network representations are not only determined by the extent (the amplitude) of rhythmic and arrhythmic activity in specific frequency bands, but also by the phase of the respective oscillations, which correlates with the net level of neural depolarization and thus reflects the excitability of nerve cells. This line of research is only in its beginnings, and many questions remain: Is there indeed a causally relevant network code in the brain, i. e. can the brain “read out” the overall activity level of these local networks, or is only the

activity of individual cells causally relevant? If there are multiple “codes”, which one is used under which circumstances? And how do network engrams relate to memory traces of individual neurons?

Fortunately, most people do not need to be implanted with electrodes. Activity of neural networks can be recorded in healthy participants as well, albeit in more indirect ways: using conventional EEG recordings attached to the scalp, and via the most important method of current cognitive neuroscience research, functional magnetic resonance imaging (fMRI). Both methods have contributed important findings to our understanding of engrams. Various fMRI studies have shown that stimulus-specific representations can be decoded from distributed BOLD activity patterns in the neocortex. Interestingly, it is much more difficult to

N. Axmacher

In search of the human engram**Abstract**

Despite abundant knowledge on the neural basis of memory functions in the human brain, stimulus-specific memory traces — engrams — have long remained elusive. In this article, recent developments that start to shed light on the mechanisms underlying the formation, modification and potential degradation of engrams are reviewed. Stimulus-specific memory representations appear to occur at different levels of brain organization, from spike rates of individual cells via time-frequency signatures of small-scale neural networks to distributed activity patterns. However, memories undergo transformation whenever they are recalled. Thus, novel methodological approaches need to be employed in order to identify considerably modified engrams. Furthermore, engrams are impaired in a number of diseases involving memory dysfunction. This article is concluded with a description of recent translational work on altered content-specific memory representations in the context of Alzheimer's disease.

Keywords

Engram · Decoding · Generative memory · Alzheimer's disease

decode content-specific representations from the hippocampus. This difficulty may be explained by results that suggest that hippocampal representations are not topologically organized — in other words, that adjacent hippocampal neurons represent very different contents. As a result, the activity of different hippocampal fMRI voxels, each of which sample the activity of a large number of neurons, only differs marginally. In addition to decoding specific representations during encoding, fMRI studies have also contributed to understanding the “fate” of memory traces when they are stabilized during post-learning consolidation stages: For example, it could be shown that stimulus-specific representations re-occur spontaneously during the awake resting state and sleep following learning, and that this reactivation is beneficial for subsequent memory ([3]; [8]; ■ Fig. 4). This process is functionally similar to the well-described reactivation of *place cells* in rodents, but occurs on a much more extended temporal and spatial scale.

Of particular importance for future research are simultaneous EEG/fMRI recordings which allow one to investigate the relationship between hippocampal activity — that can be indirectly measured via fMRI — and stimulus-specific representations in EEG oscillatory patterns. Such experiments allow one to assess how the hippocampus, or other areas like the prefrontal cortex, control the encoding of stimulus-specific representations and their transformation into more enduring engrams. Moreover, recordings at increasingly higher magnetic field strengths allow the activity of smaller brain regions down to individual cell layers to be investigated, which enables one to test more mechanistic predictions derived, for example, from animal experiments, intracranial EEG recordings, or computer simulations.

Storage versus transformation

Memory is not just a passive storage device. In some cases it can be important to retrieve information exactly as it has been previously learned (for example, if I forgot my shopping list and try to remember which groceries I need). In many other

situations, however, memory is selective and generative, and for good reasons. It selectively allows me to remember exactly those aspects of an episode that are relevant in a given situation and to suppress others. I can either concentrate on specific details (where exactly did I park my car yesterday?) or on general facts (what size does a parking space need before my car to fit it?). I may try to remember unpleasant or conflict-related interactions as little as possible — but if I do, then I may transform these memories so that my role in a conflict does not appear too negative. Our memory is thus continuously reconstructing our life stories. For example, a number of famous experiments by Elizabeth Loftus showed that participants can be made to believe, via suggestive questioning, that they have experienced specific episodes that actually did not occur. Similar modifications presumably also occur outside of the laboratory, suggesting that our view of our past — our autobiographical memory — is indeed rather flexible.

A number of neuroscientific studies in rodents, as well as psychological experiments in human subjects, have provided converging evidence that memory traces become labile when they are recalled. Afterwards, they may be either deleted, restored in a more or less modified form, or undergo a process of renewed stabilization called reconsolidation. Permanently fixed and rigid memory traces do not seem to be the normal case, but rather occur in psychiatric conditions. In particular, patients suffering from posttraumatic stress disorder re-experience the same traumatic episodes repeatedly and without alterations; they cannot distance themselves from these experiences or integrate them into their autobiographical memory.

Within the cultural sciences, the dichotomy between a storage and a constructivist or generative model of memory is well established [1]. Within the cognitive neurosciences, however, a unidimensional concept of memory as a storage device still prevails [2]. An important future research direction thus consists of investigating both the identical reproduction and the modified reconstruction of engrams. When and how

are engrams modified after their initial formation? Which brain areas support the identical reproduction of a certain memory trace, which favor its modification, and which may be important for deleting it? How can painful experiences be transformed — for example via psychotherapy — such that they can be eventually integrated into our self-image? These questions are not only relevant for basic research on memory, but could also lead to novel therapies in diseases such as posttraumatic stress disorder. On a methodological level, novel approaches such as the use of *forward encoding* models are particularly relevant here: Based on extensive fMRI recordings, explicit receptive field models of individual voxels can be estimated that allow for a direct visualization of neural representations of novel stimuli. In contrast to conventional

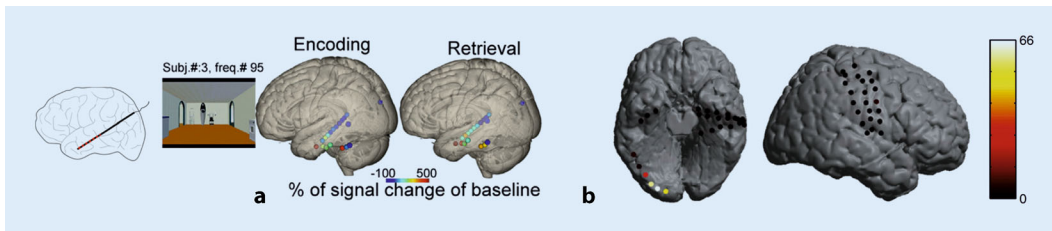


Fig. 3 ▲ Stimulus-specific representation in networks (intracranial EEG). Intracranial EEG recordings show stimulus-specific activity patterns during navigation in a virtual environment (a) or during presentation of different letters (b). a Snap shot of activity during navigation through a virtual room and color-coded activity distribution in the brain during learning (*Encoding*) and recall (*Retrieval*) of the path through this room. b Color-coded contribution of activity in single electrodes to representations of specific letters. Figures reproduced with permission from [10] (a) and [9] (b)

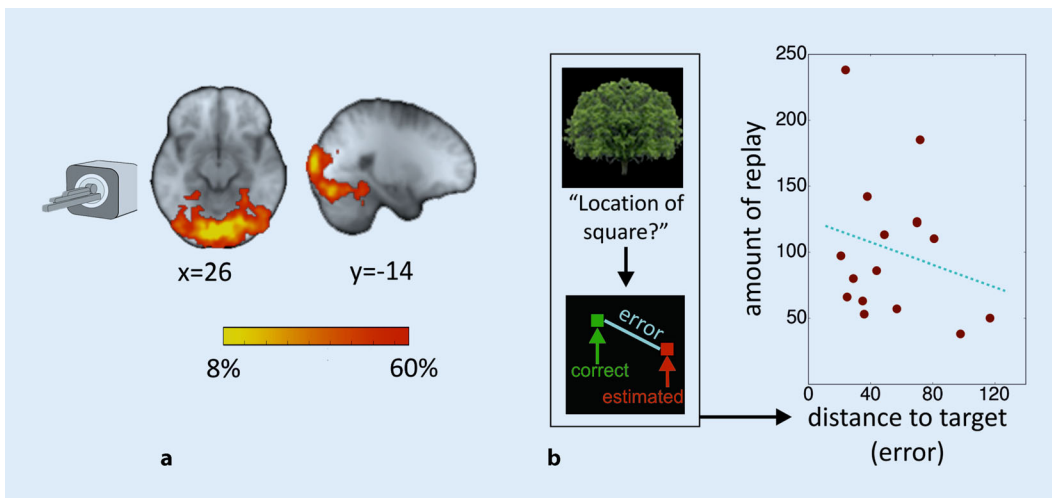


Fig. 4 ▲ Stimulus-specific representation in networks (functional magnetic resonance imaging, *fMRI*). Recordings via *fMRI* show that the reactivation of specific memory traces during quiet resting state and sleep is beneficial for memory consolidation. a Relevant brain areas. b Intra-individual correlation between the number of reactivations of stimulus-specific memory traces and subsequent memory for the corresponding contents. Memory traces that were reactivated more often (higher “amount of replay”) were afterwards remembered more accurately (smaller “distance to target”). Figure by Lorena Deuker

MVPA approaches, forward models thus do not only enable measuring an identical reactivation of engrams, but also assessing their specific transformation — for example, during memory recall, in specific social interactions, or as a result of a psychotherapeutic or drug intervention.

The impaired engram

Memory disorders are, after attention deficits, the second most frequent neuropsychological symptom and the cardinal symptom of Alzheimer’s disease (AD). On a neuropathological level, AD is associated with the formation of extracellular amyloid plaques and intracellular deposits of hyperphosphorylated tau proteins. Among the brain regions affected earliest in the disease are areas

that play an important role for memory functions. Not all memory processes are equally impaired in AD patients; instead, they have particularly early deficits in precisely differentiating between relatively similar events. This suggests that their ability to build and maintain specific engrams is reduced.

Currently there is no curative treatment for Alzheimer’s disease, despite several very large and costly studies using novel therapeutic drugs (in particular, involving several kinds of antibodies). These studies have only resulted in relatively minor effects. Does this mean that they were based on incorrect pathophysiological assumptions — that the mechanism of action of these drugs was badly chosen? Not necessarily; it is more likely that therapeutic attempts started too late, when large brain areas

were already affected by Alzheimer’s pathology and many neurons had already been destroyed. Therefore, one major focus of ongoing Alzheimer’s disease research consists in the investigation of early disease stages, in the identification of risk factors and the development of novel biomarkers. The earliest possible risk factor for a disease consists in genetic polymorphisms, which are associated with a higher incidence of this disease. The common form of Alzheimer’s disease — the one starting relatively late in life — is not monogenetically determined. However, one particular genetic risk factor is primarily associated with an increased disease risk: the epsilon 4-allele of the apolipoprotein E gene. Homozygous Apo-E4 carriers have a more than tenfold increased risk for developing Alzheimer’s dementia,

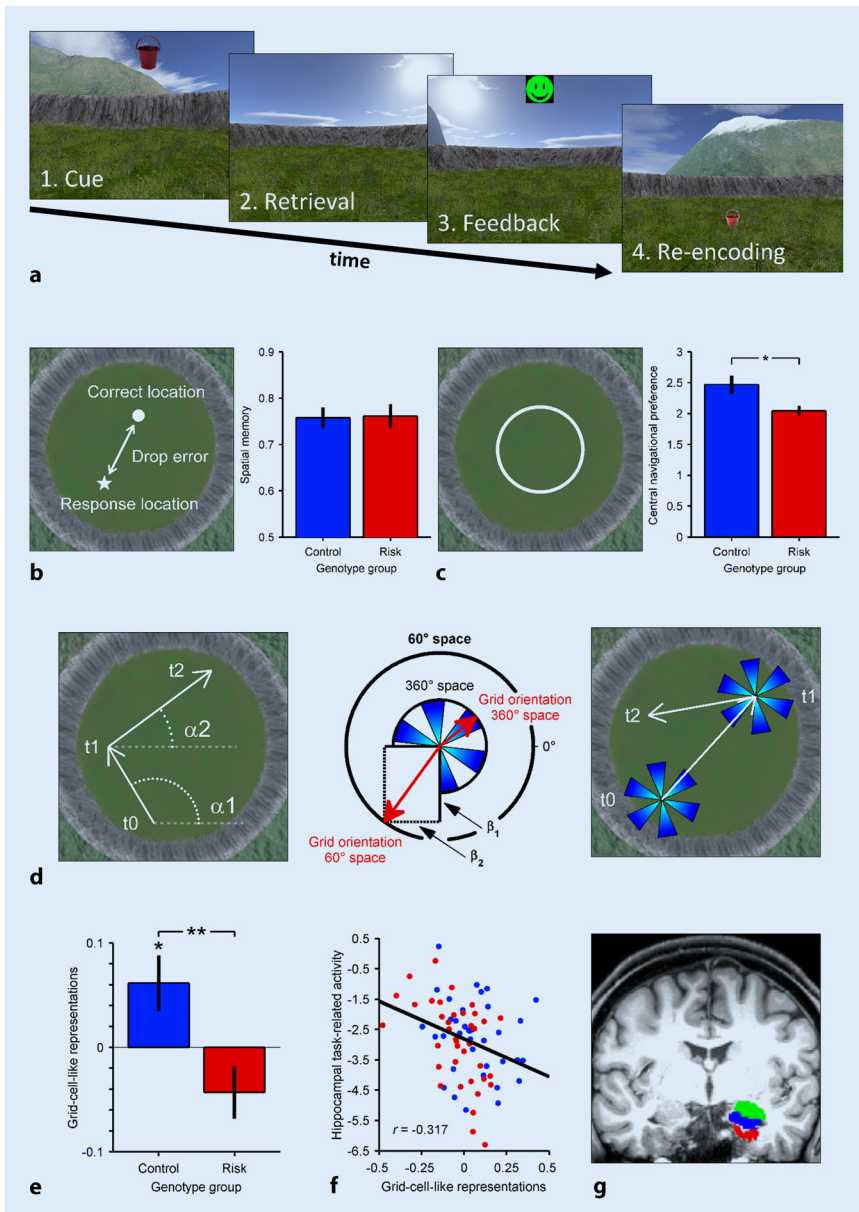


Fig. 5 ▲ Impaired grid cell representations in the brain of genetic Alzheimer's disease risk carriers. **a** Virtual navigation task in the MRI scanner. Participants have to remember the location of different objects in a virtual arena and to drop these objects at the correct position. **b** Genetic risk carriers (red) do not have higher "drop errors" than control participants (blue), indicating that they do not show any overt impairment in spatial memory. **c** Risk carriers show altered navigational behavior: They navigate less often at the center of a virtual arena. **d** Analysis strategy for detection of *grid cell* representations in fMRI data: At each moment in time during spatial navigation, the direction of movement in the arena is recorded, and preferred directions of *grid cell* representations are extracted in one half of the data. These axes are sixfold rotationally symmetric (blue bars of the circle in middle and right subpanels). In the other half of the data, fMRI activity during movements aligned or misaligned with respect to these preferred directions is contrasted. **e** Magnitude of *grid cell* representations is significantly impaired in genetic risk carriers as compared to control subjects. **f** Compensatory increase in hippocampal task-related activity in participants with reduced *grid cell* representations. **g** Investigated brain area (entorhinal cortex: red) and adjacent areas (hippocampus, blue; amygdala, green). Figure by Lukas Kunz

but the homozygous occurrence of this allele is relatively rare (1 in 100). However, heterozygous carriers are much more frequent — one person out of six carries one Apo-E4 allele, and their risk is already increased by a factor of three.

The entorhinal cortex is among the first regions affected by Alzheimer's pathology and plays an important role for spatial navigation and memory. The 2014 Noble Prize in Physiology or Medicine was awarded to the discovery of *grid cells* in the entorhinal cortex, which show a specific sixfold rotationally symmetric activity pattern during spatial navigation. Interestingly, the spatial axes of this pattern are not distributed randomly across all *grid cells* in the entorhinal cortex, but are relatively uniform. It is therefore possible to detect their sixfold rotationally symmetric representational pattern even using macroscopic activity via fMRI and to compare the magnitude of this pattern between different subject populations. In a recently published study, we could show that the relatively frequent heterozygous Apo-E4 carriers show a marked impairment in their *grid cell* representations already at a surprisingly young age, in their early twenties ([5]; ■ Fig. 5). In addition, their navigational behavior was significantly altered, because they were navigating more often at the borders of a virtual arena than control participants without any Apo-E4 allele. No impairments in spatial memory were found in subjects with reduced *grid cell* representations; this is probably due to a compensatory hyperactivity in the hippocampus. This study measured *grid cell* representations only indirectly via fMRI; follow-up studies thus have to investigate if and how these network representations are related to the activity of single *grid cells*. From a clinical perspective, it is also important to determine whether alterations of *grid cell* representations and of engrams of spatial positions are indeed related to early Alzheimer's pathology, and if they may serve as novel biomarkers for the detection of Alzheimer's dementia in older adults at a stage when the disease may still be treatable.

Outlook

The application of increasingly advanced analysis methods is bringing us closer to identifying content-specific memory traces — engrams — in single cells, local EEG rhythms, and distributed fMRI patterns of the human brain. These developments are contributing to a mechanistic understanding of how specific experiences are encoded into memory traces, how they are subsequently transformed, and how they are affected by neurological and psychiatric diseases. In the future this will lead to a better understanding of why some memories remain permanently stable, while others change during life and still others undergo complex distortions. Neuroscientific research may thus also contribute to answering questions that were traditionally asked in the humanities: How do we perceive the world? How do we remember our past? How does our identity emerge from our memories?

Corresponding address



Prof. Dr. N. Axmacher

Abteilung für Neuropsychologie, Institut für Kognitive Neurowissenschaft, Fakultät für Psychologie, Ruhr-Universität Bochum
Universitätsstraße 150,
44801 Bochum, Germany
nikolai.axmacher@rub.de

Prof. Dr. Nikolai Axmacher studied medicine and philosophy in Berlin and Paris. Following a dissertation at the Institute of Physiology and the Institute of Theoretical Biology in Berlin, he moved to the Department of Epileptology at the University of Bonn, where he completed his habilitation in Cognitive Neuroscience in 2009 and became an Emmy Noether research group leader in 2011. The same year, he also became a junior research group leader at the German Centre for Neurodegenerative Diseases in Bonn. Since 2014 he is head of the Department of Neuropsychology, Faculty of Psychology, Ruhr University Bochum. Axmacher is working on the neural mechanisms of memory processes (short-term memory maintenance, long-term memory formation, memory consolidation, autobiographical memory) and memory dysfunctions (due to repression of intrapsychical conflicts, in posttraumatic stress disorder and related to Alzheimer's disease). He combines fMRI, EEG and simultaneous EEG/fMRI recordings in healthy participants and Alzheimer's disease risk carriers with intracranial EEG recordings in neurological and psychiatric patients.

Compliance with ethical guidelines

Conflict of interest. N. Axmacher states that there are no conflicts of interest.

The accompanying manuscript does not include studies on humans or animals.

References

1. Assmann A (1999) Erinnerungsräume. Formen und Wandel des kulturellen Gedächtnisses. C.H. Beck, München
2. Axmacher N, Do LAT, Kessler H, Fell J (2010) Natural memory beyond the storage model: repression, trauma, and the construction of a personal past. *Front Hum Neurosci* 4:211
3. Deuker L, Olligs J, Fell J, Kranz TA, Mormann F, Montag C, Reuter M, Elger CE, Axmacher N (2013) Memory consolidation by replay of stimulus-specific neural activity. *J Neurosci* 33(49):19373–19383
4. Ekstrom AD, Kahana MJ, Caplan JB, Fields TA, Isham EA, Newman EL, Fried I (2003) Cellular networks underlying human spatial navigation. *Nature* 425(6954):184–188
5. Kunz L, Schröder TN, Lee H, Montag C, Lachmann B, Sariyska R, Reuter M, Stirnberg R, Stöcker T, Messing-Floeter PC, Fell J, Doeller CF, Axmacher N (2015) Reduced grid-cell-like representations in adults at genetic risk for Alzheimer's disease. *Science* 350(6259):430–433
6. Lashley KS (1950) In search of the engram. *Symp Soc Exp Biol* 4:454–482
7. Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I (2005) Invariant visual representation by single neurons in the human brain. *Nature* 435(7045):1102–1107
8. Staresina BP, Alink A, Kriegeskorte N, Henson RN (2013) Awake reactivation predicts memory in humans. *Proc Natl Acad Sci USA* 110(52):21159–21164
9. van Gerven MA, Maris E, Sperling M, Sharan A, Litt B, Anderson C, Baltuch G, Jacobs J (2013) Decoding the memorization of individual stimuli with direct human brain recordings. *Neuroimage* 70:223–232
10. Zhang H, Fell J, Staresina BP, Weber B, Elger CE, Axmacher N (2015) Gamma power reductions accompany stimulus-specific representations of dynamic events. *Curr Biol* 25(5):635–640