Recipe for a conspiracy theory. Memes and neuroscience.

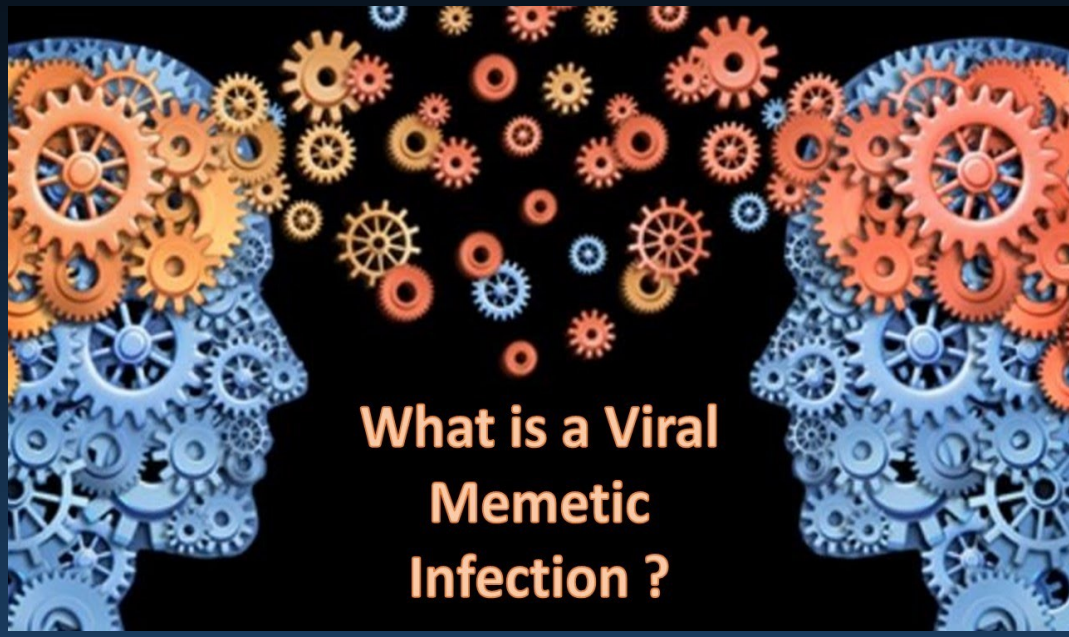
Włodzisław Duch

Neurocognitive Laboratory, Center for Modern Interdisciplinary Technologies,
& Dept. of Informatics, Faculty of Physics, Astronomy & Informatics,
Nicolaus Copernicus University

Google: Wlodzislaw Duch

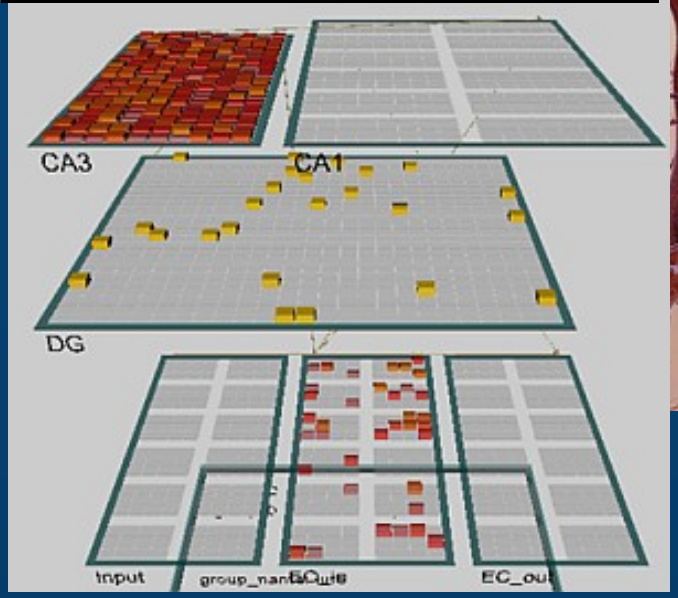
WFAiIS UMK 10/3/2022

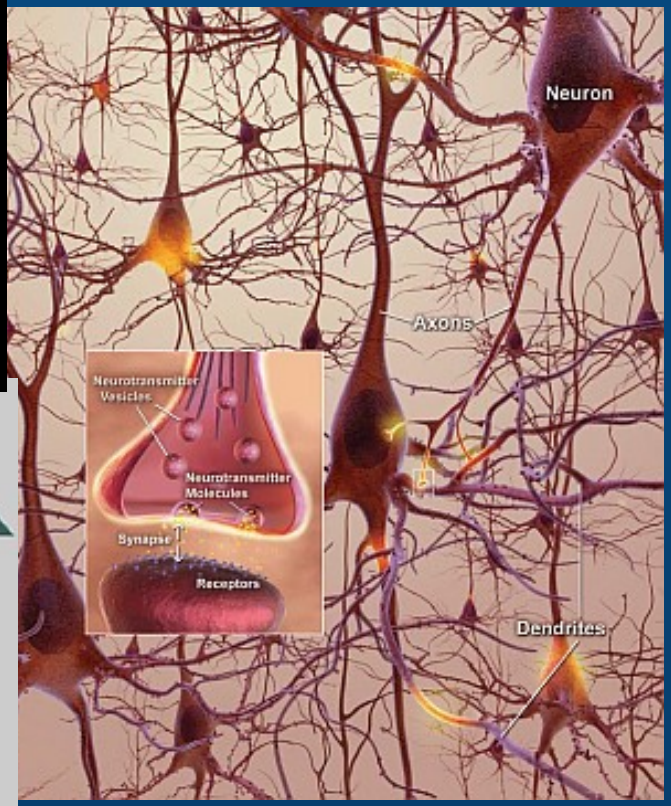
- Phenomics and understanding of human behavior.
- Memetics and psychological perspective.
- Brain states and memory.
- Hippocampus and memory formation.
- We need to simplify …
- Conspiracies in our brains.
- Direct brain-brain? Perspectives.



Duch W. (2021). *Memetics and Neural Models of Conspiracy Theories*. Patterns 2(11), 2-13. More papers on these topics.

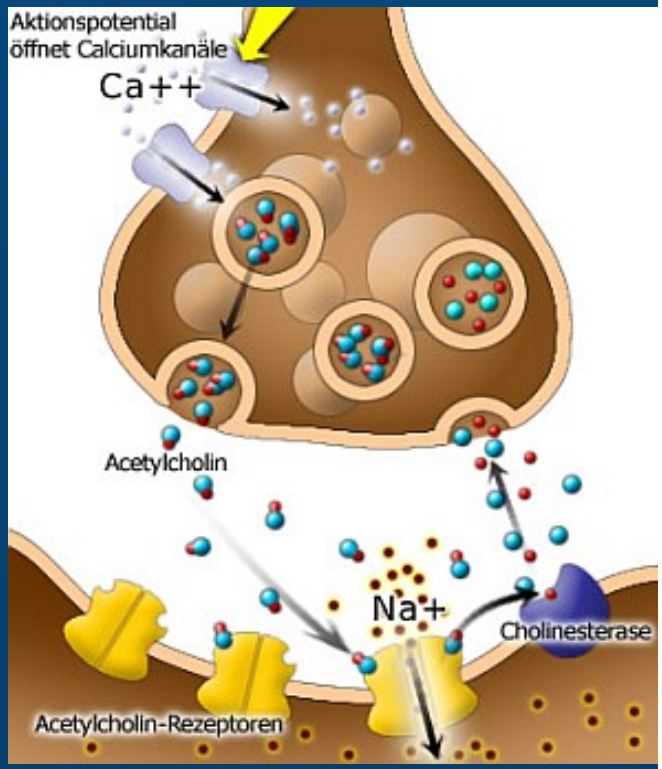
From Behavior to Neurons

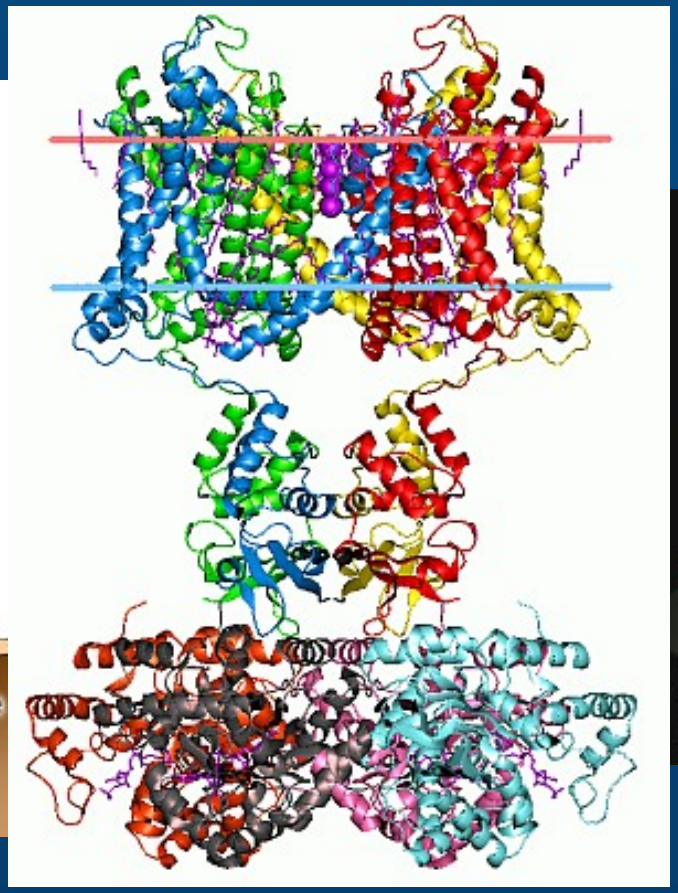


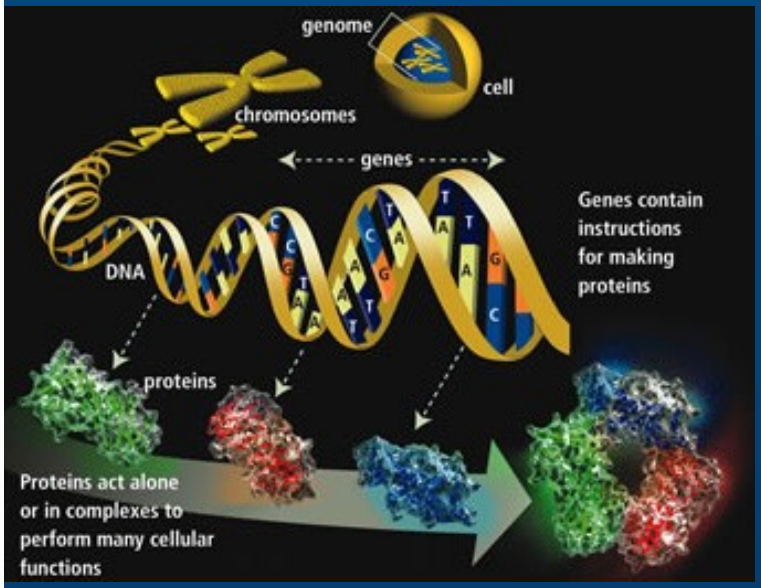


Cognitive phenotypes ↔ behavior ↔ conspiracies ↔ memory, memes
↔ **neurodynamics** ↔network structure ↔ connectome ↔ neuron properties

From Neurons to Genes

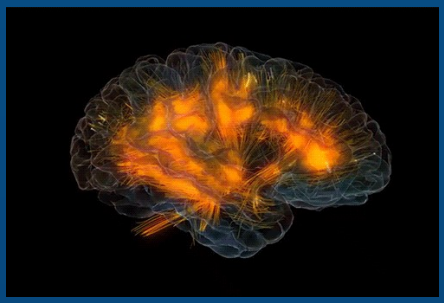






↔ **neurodynamics** ↔network structure ↔ connectome ↔ neuron properties ↔ synapses ↔ receptors, ion channels ↔ neurochemicals, proteins ↔ genes.

Explaining behavior



Seconds

Minutes

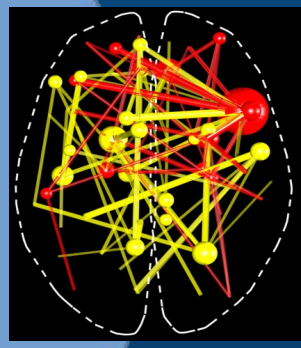
Days

Month

Years

Millenia

Eras/Eons



Cognitive phenomics, fast/slow: neurodynamics, hormones, education, culture, infancy, gestation and evolution.

Phenomics => behavior

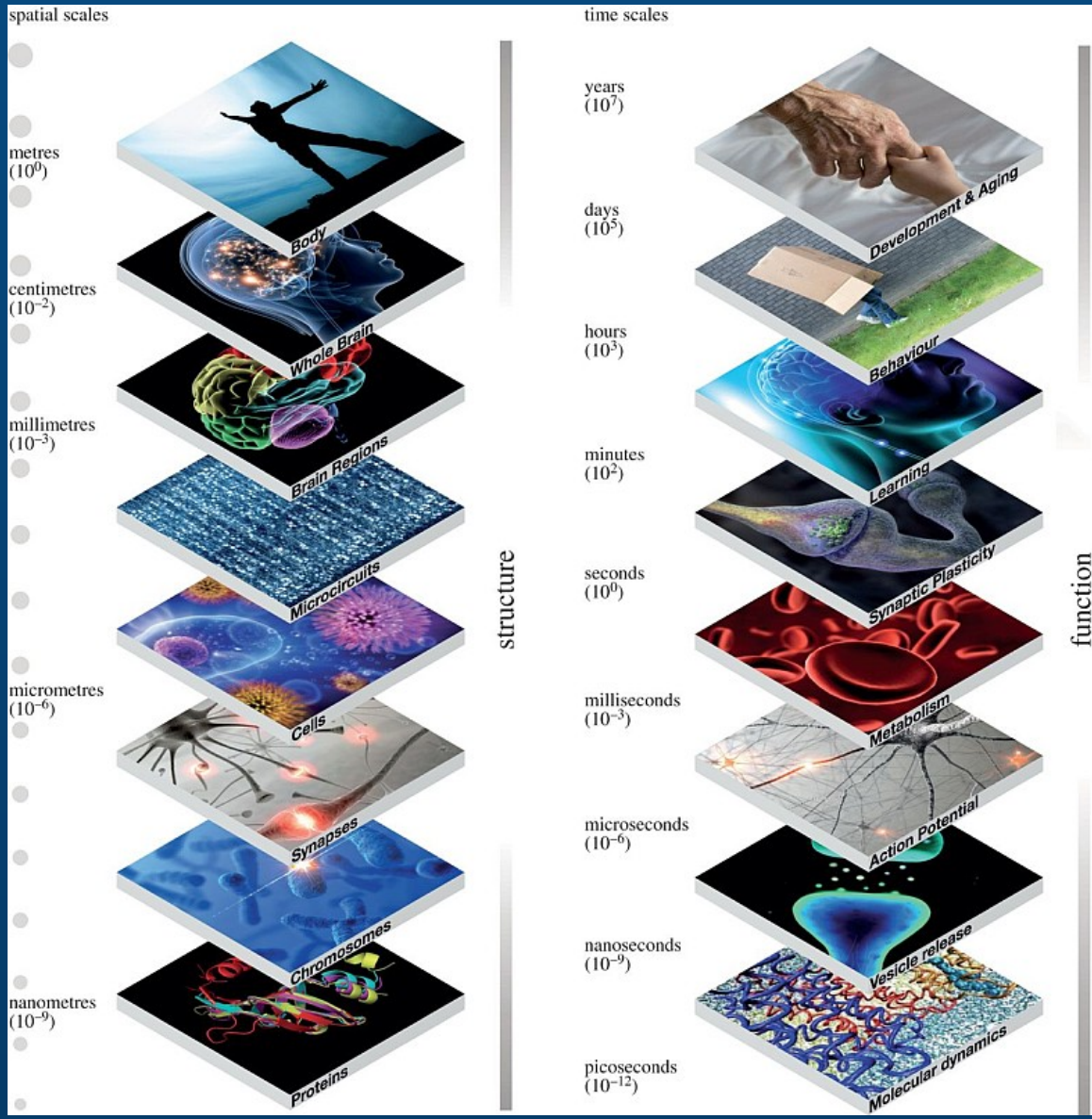
The Consortium for Neuropsychiatric Phenomics (2008).

"… categories, based upon presenting signs and symptoms, may not capture fundamental underlying mechanisms of dysfunction" (Insel et al., 2010).

New approach: RDOC NIMH.

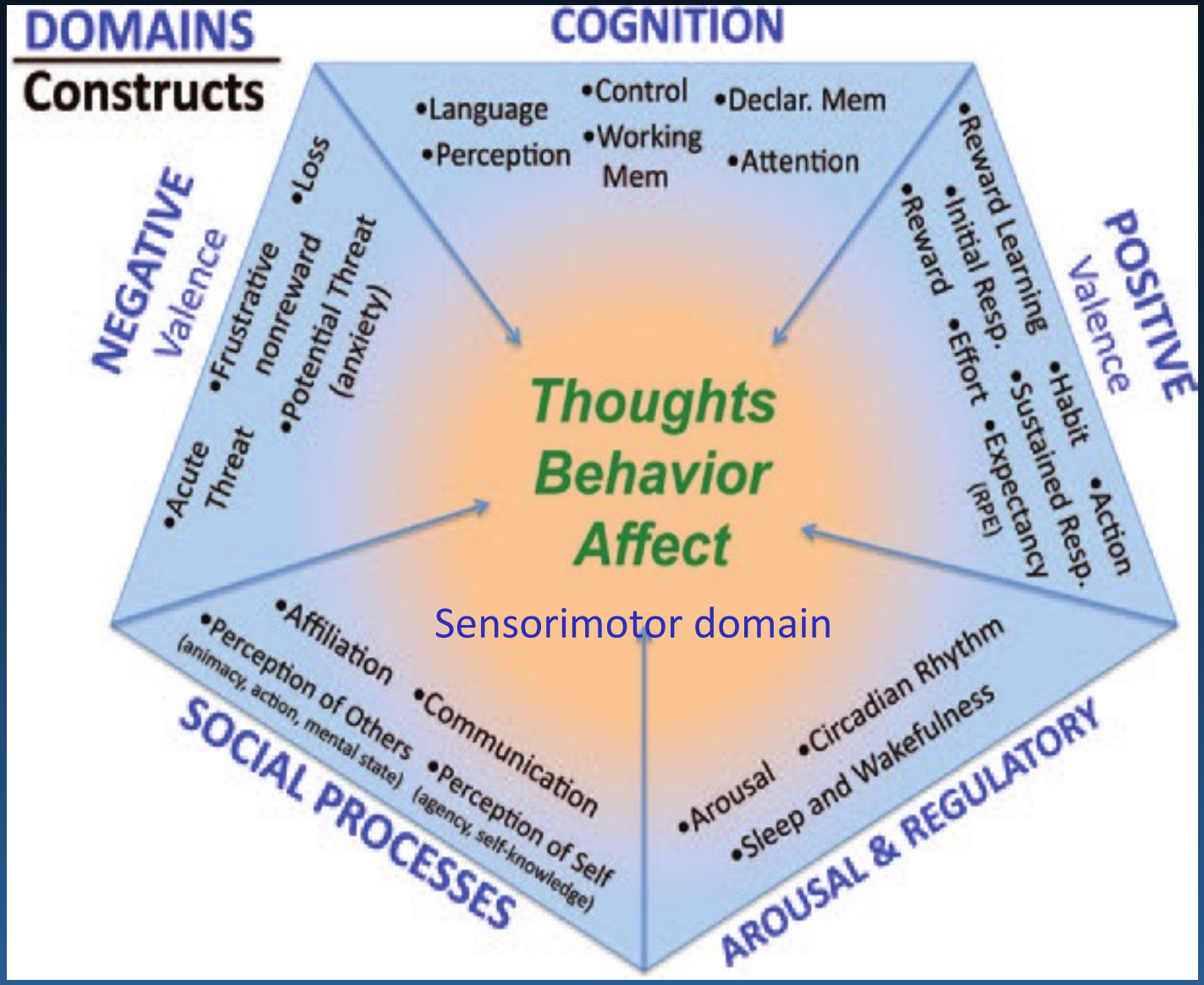
Description of organisms at different levels will help to answer different types of questions.

Network level is in the middle and can be connected to the mental level via computational models.



RDoC
Research Domain Criteria Initiative

NIMH RDoC Matrix for deregulation of 6 large brain systems.



Psychological perspective

Memetics

Richard Dawkins introduced memetics in 1976 in "The Selfish Gene" book. Evolution requires three processes: inheritance, variation, and selection pressure.

Meme: an information unit, cultural idea, symbol, or practice that can be transmitted from one mind to another through writing, speech, gestures, rituals, imitation. Memes like genes, are inherited (communicated), undergo variations & rapid selection.

Memeplex – is a collection or grouping of memes that have evolved into a mutually supportive or symbiotic relationship, a set of ideas that reinforce each other.

An example of a memeplex: a political ideology, subculture, sport, religion.

Meme pool – a population of interbreeding memes.

Memetic engineering – deliberately creating memes using engineering principles.

Memetic algorithms – an approach to evolutionary computation that attempts to emulate cultural evolution in order to solve optimization problems.

Memeoid – person completely overtaken by memeplex, to the extent that their own survival becomes inconsequential. Ex. kamikazes, suicide bombers, cult followers.

Memes have been proposed as a scientific discipline in cultural studies, and as the center of a theory of mind. **But there is nothing rational about memes.**

The Cognitive Bias Codex

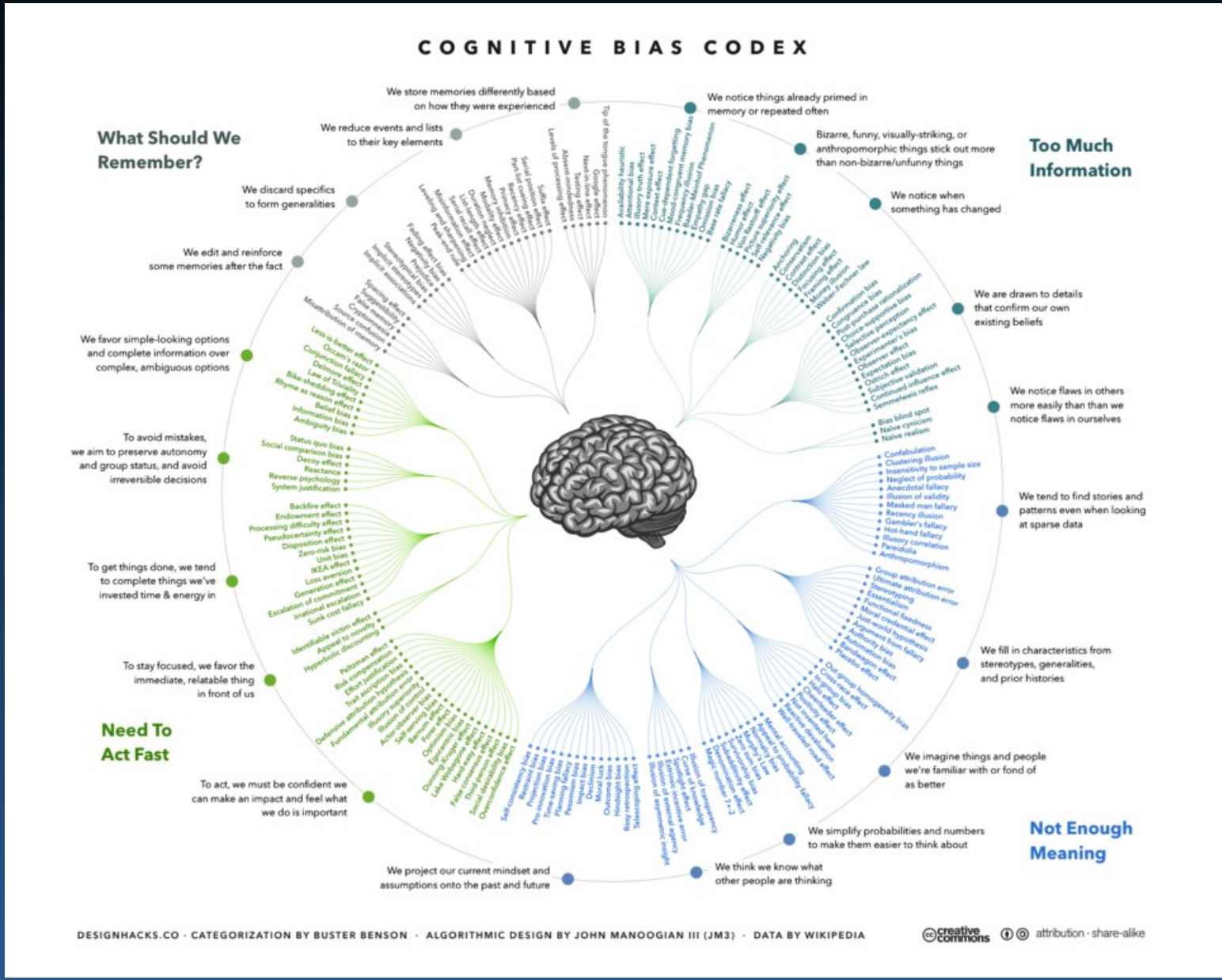
Cognitive biases are systematic patterns of deviation from rationality in judgment, studied in psychology, sociology, behavioral economics.

Over 180 biases …

Nobel 2002 in Economic Sciences

Daniel Kahneman, Bounded Rationality

Nobel 2017, **Richard H. Thaler**, decision making.



Source: Wikimedia commons

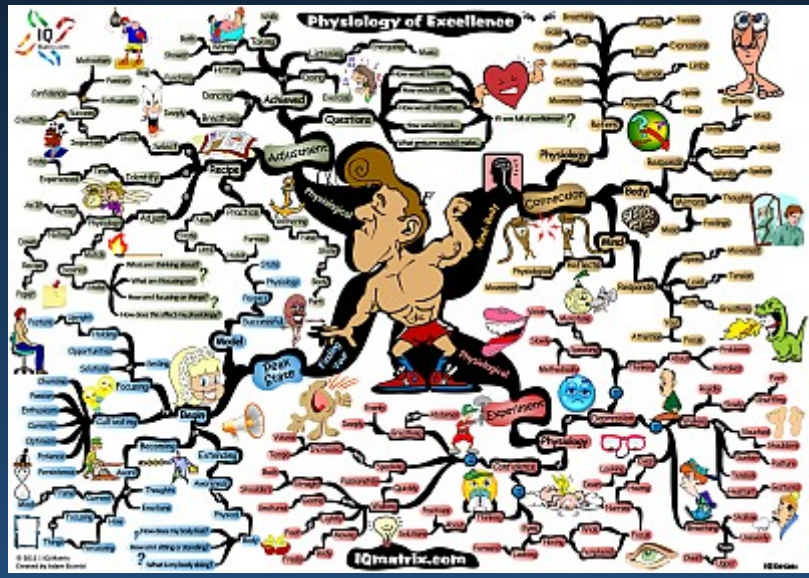
Conceptual space

A shared conceptual space is necessary for communication, for understanding, reading and writing with comprehension.

Since 1986, the Core Knowledge Foundation has been working in the US and UK trying to define this common cultural code, from kindergarten to the end of elementary school.

Communication space: how we understand a concept, what associations we have, what chain of associations, what resonance does it have in another brain?

What kind of conceptual grid do we have?



Natural language models, such as OpenAI GPT-3, or Google Switch Transformer (1600 billion parameters, in 2048 domains) are trained on general knowledge, trillions of words, and then used to create models in a specialized domain.

Words, concepts that activate the brain enable segmentation of experience, without them there would be a constant flow of states, very limited planning, poor neuronal space, in which associative processes take place.

Conceptual framework, "mind space" (WD, 1994).

IQmatrix.pl and mind maps.

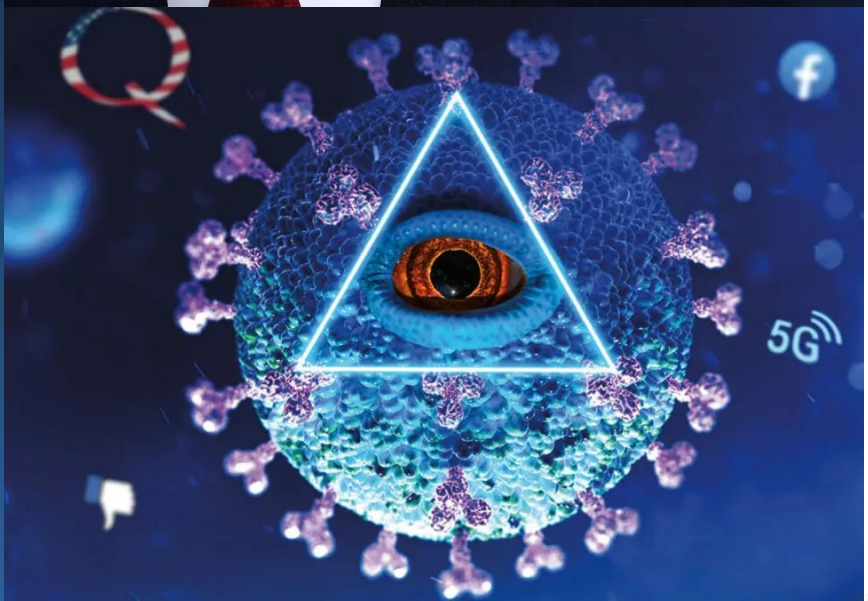
5G
DOES NOT
SPREAD
COVID-19
STOP
5G

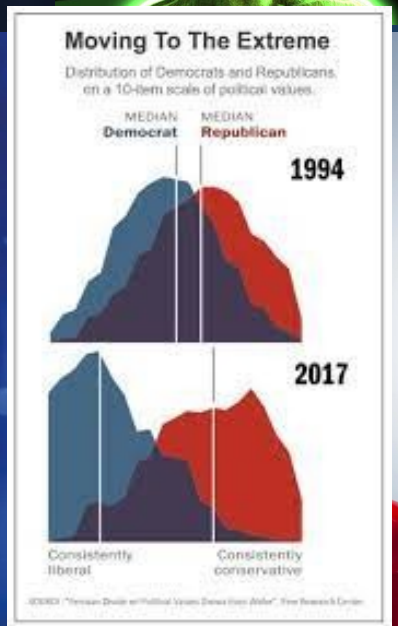
INVESTIGATE
COVID
AND 5G!

I WANT TO
WE COME IN PEACE
Believe

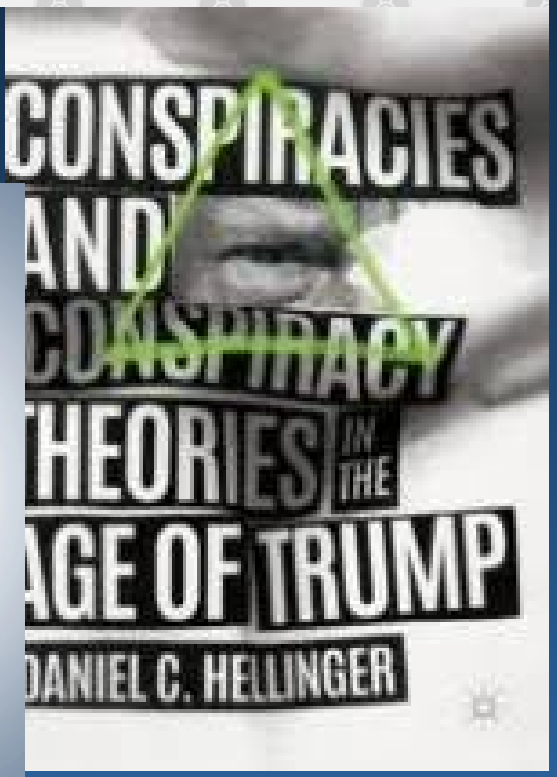
CNN
CNN SPECIAL REPORT
DONALD TRUMP'S
CONSPIRACY
THEORIES
MONDAY 9P ET/PT

THE TRUTH
BEHIND
AREA 51

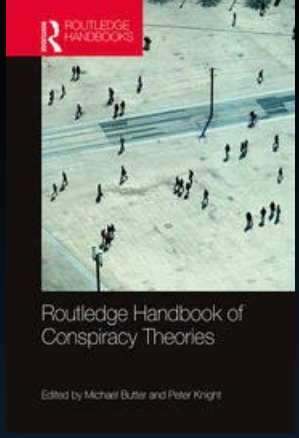

5G


Moving To The Extreme
Distribution of Democrats and Republicans on a 10-item scale of political values
MEDIAN Democrat
MEDIAN Republican
1994
2017
Consistently liberal
Consistently conservative

COVID-19


CONSPIRACIES
AND
CONSPIRACY
THEORIES IN THE
AGE OF TRUMP
DANIEL C. HELLINGER

Research on conspiracy theories



EU COST network on “Comparative Analysis of Conspiracy Theories” (COMPACT) gathered researchers in history, sociology, psychology and political sciences to “generate the thorough comprehension of the history, politics, sociology, rhetoric and psychology of conspiracy theories needed to counter their often harmful effects on democratic values”.

Results: Routledge Handbook of Conspiracy Theories (2020), 800 pages, book series on conspiracy theories - all completely ignoring neurobiological mechanisms, Infographics.

Artificial neural networks: focused on the best learning methods, not memory errors.

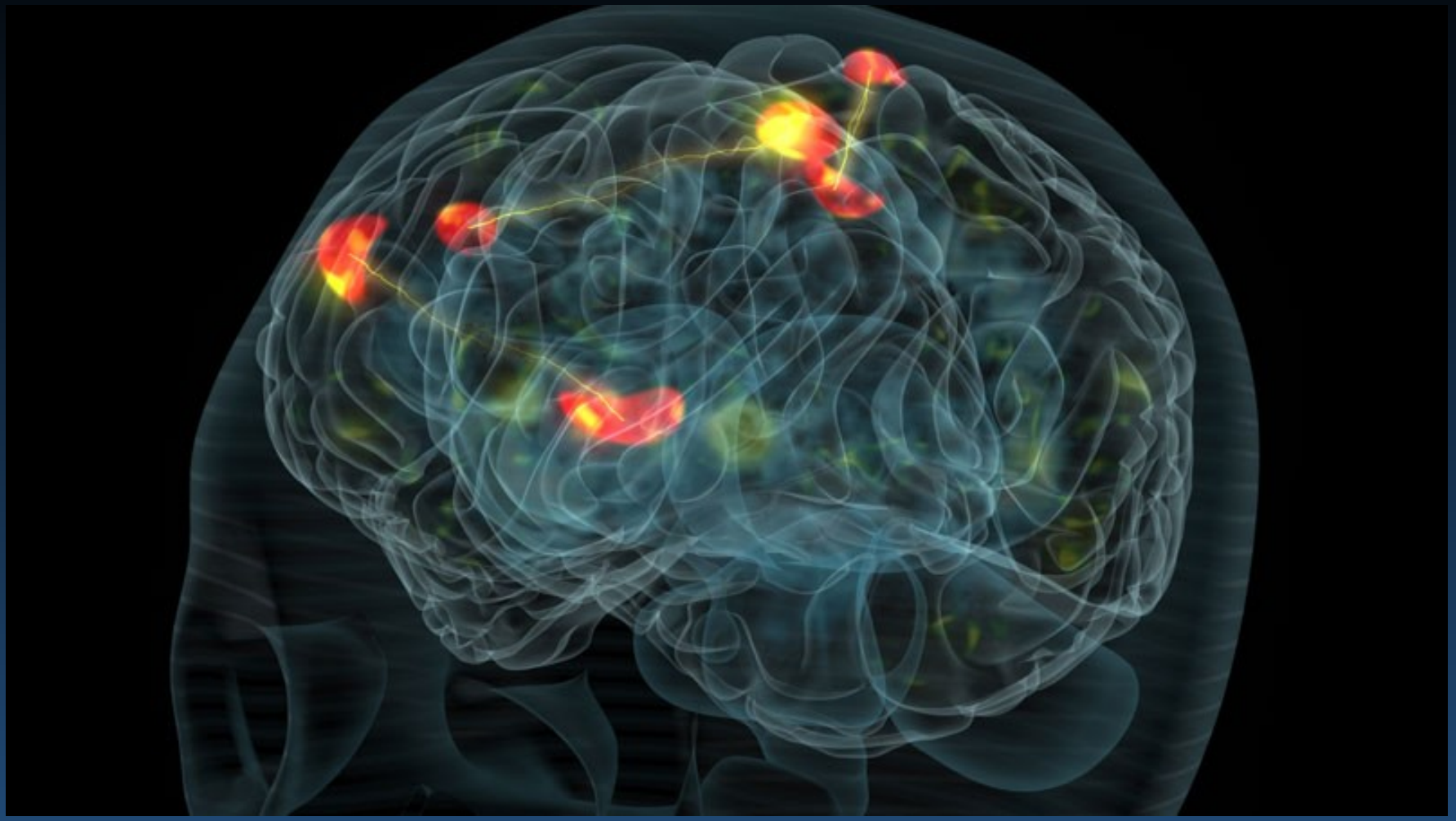
I like what I know and I know what I like …
Consolidation of new memory states in the neocortex may occur quite quickly if they form strong connections to other memory states – our conceptual networks.
Tse et al. (2007) Schemas and Memory Consolidation. Science, 316

Neural models of schemas and sequences of associations may be based on attractor states in neural networks. Each episodic or semantic memory state is based on activations of synchronized, distributed network of brain regions.

Simple explanations are rewarding, save energy needed for thinking, create false impression of reducing uncertainty. Any reference to the false information encodes it in a stronger way in the memeplex.

Brain states and memory

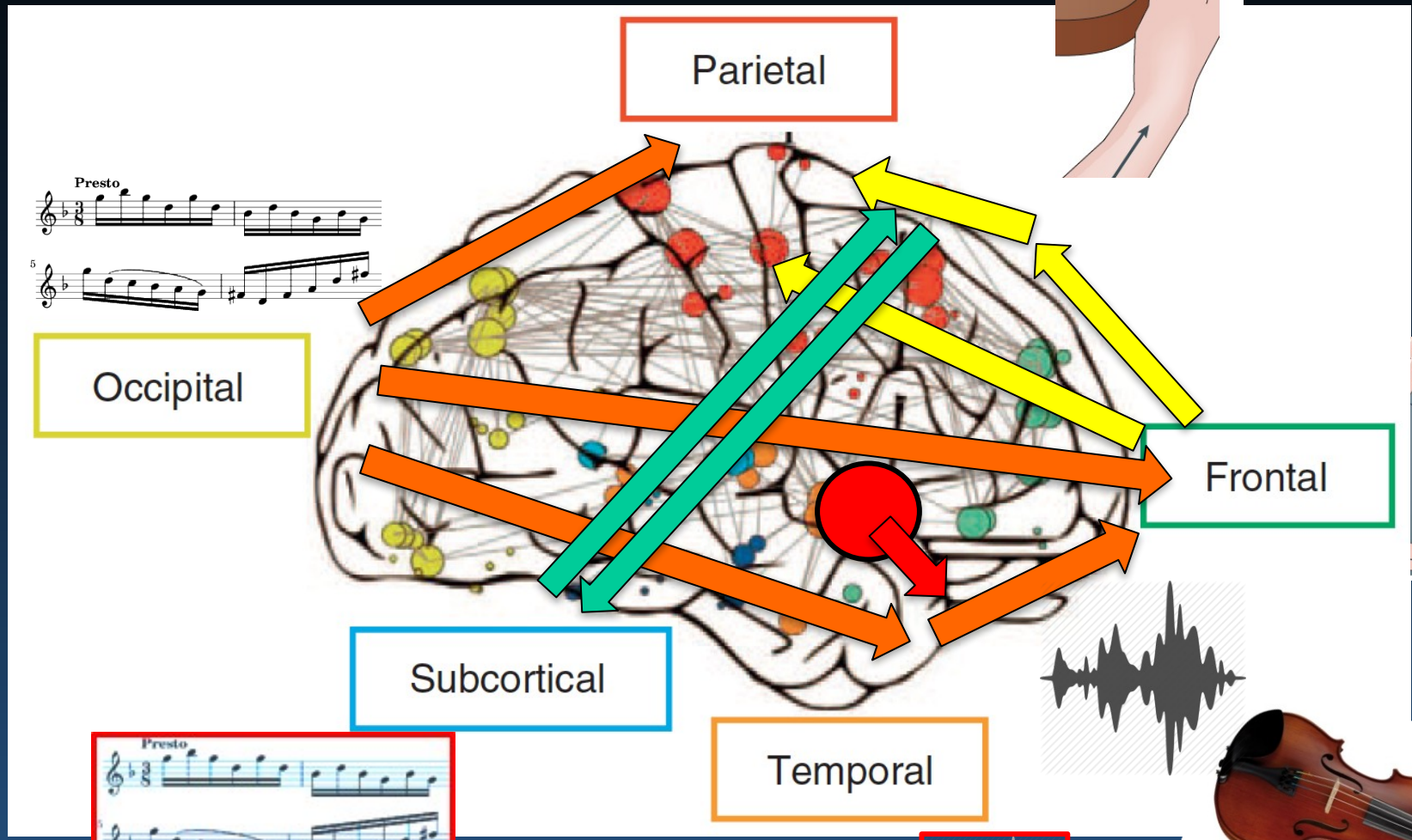
Mental states: strong, consistent activations



Large number of neuronal processes is going on in the brain at each moment.

Numerous neuronal networks excite and compete with each other.
Only the strongest processes that can be clearly distinguished from the noise (signal detection theory), will be expressed as action, speech, thought, or percept.

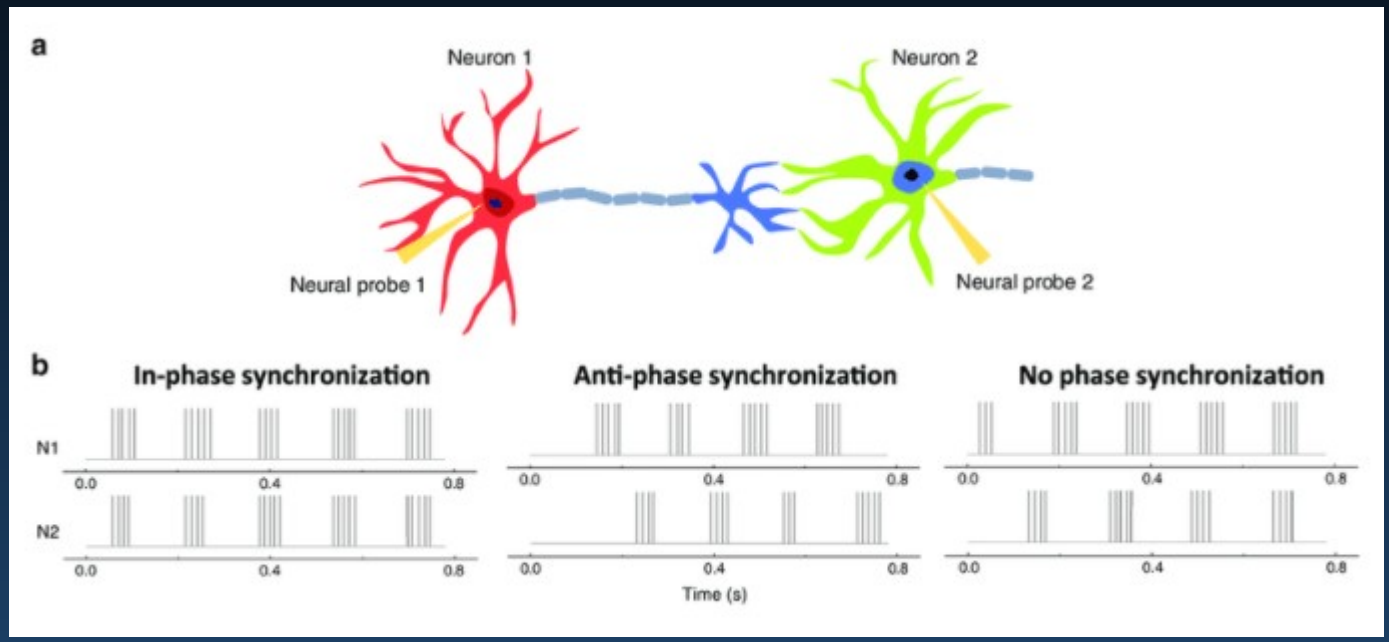
Consciousness is the perception of what is going on in one's own mind (J. Locke, 1689).



Playing an instrument engages all parts of the brain, and requires precise synchronization of neurons!

Communication between neurons

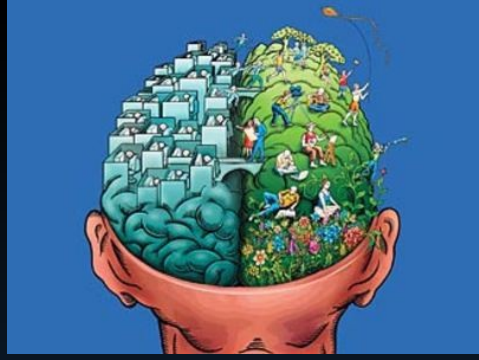
Synchronization of action potentials, changing the strength of synaptic connections (slow, LTP, LTD) but also neuronal excitability thresholds (fast).



The transmission of information at the macro level is greater than at the micro level! At the micro level it is impossible to combine different types of information with each other, reduction to the micro level is therefore a loss of information. Transmission using words is better than sending neural spikes!

Hoel, E. P. (2017). When the Map Is Better Than the Territory. Entropy, 19(5), 188

Brains and memes



Memes are “units of cultural information” (Dawkins, 1976).

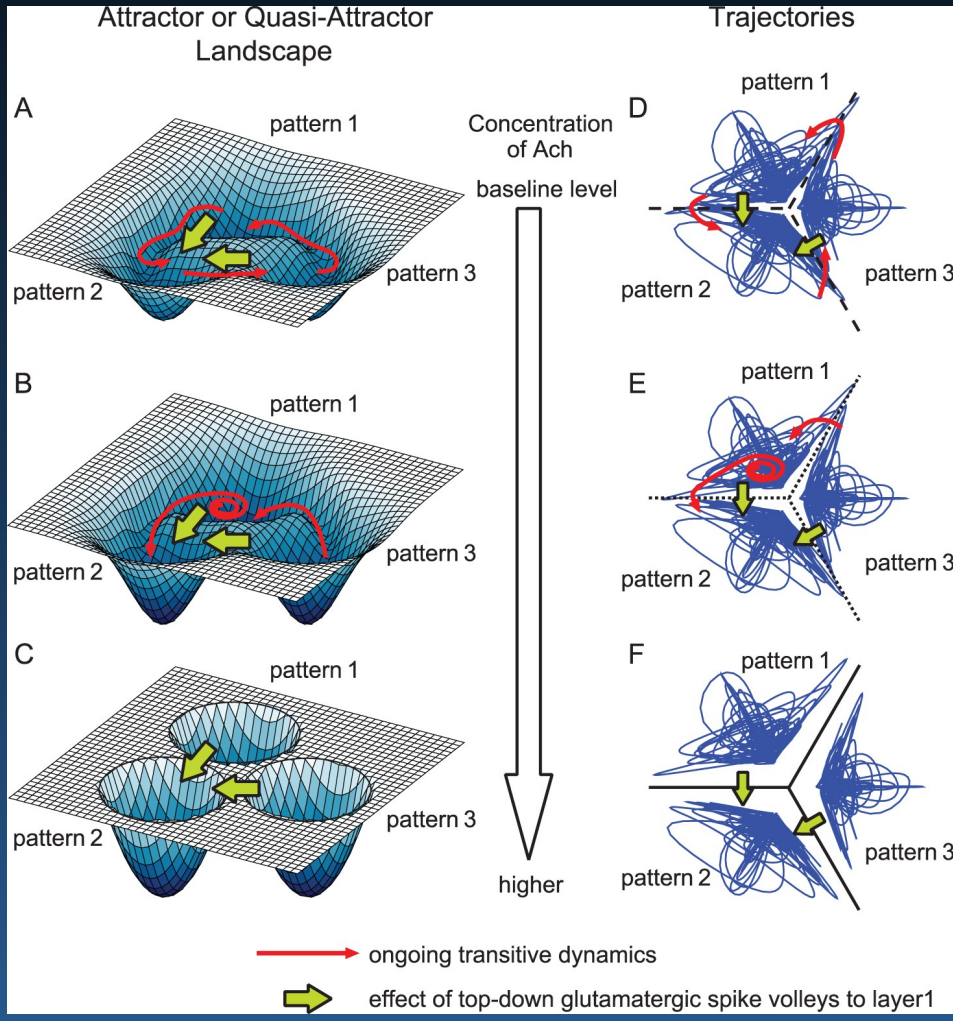
But … what is the physical carrier of memes? Neural foundations of memetics have not been developed.

Memes are memory patterns much less stable then genes. Brain is a dynamical system!

Memes are attractors of neurodynamics. Quasi-stable activations of the brain.

Amit, D. J. (1992). Modeling Brain Function: The World of Attractor Neural Networks. CUP.

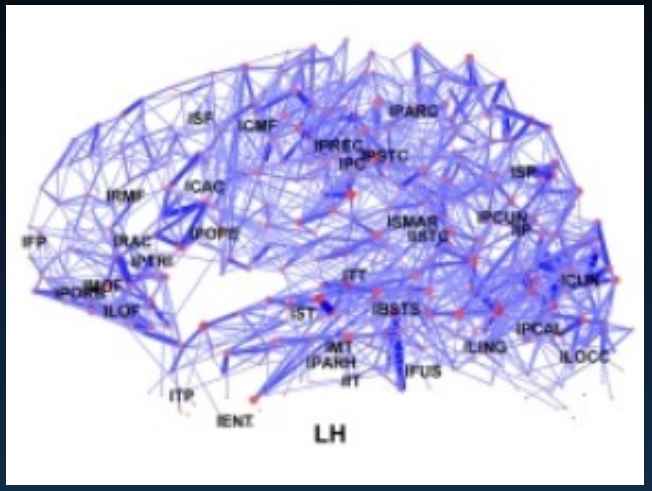
Kanamaru T, Fujii H & Aihara K. (2013). Deformation of Attractor Landscape via Cholinergic Presynaptic Modulations: A Computational Study Using a Phase Neuron Model. PLOS ONE, 8(1), e53854.

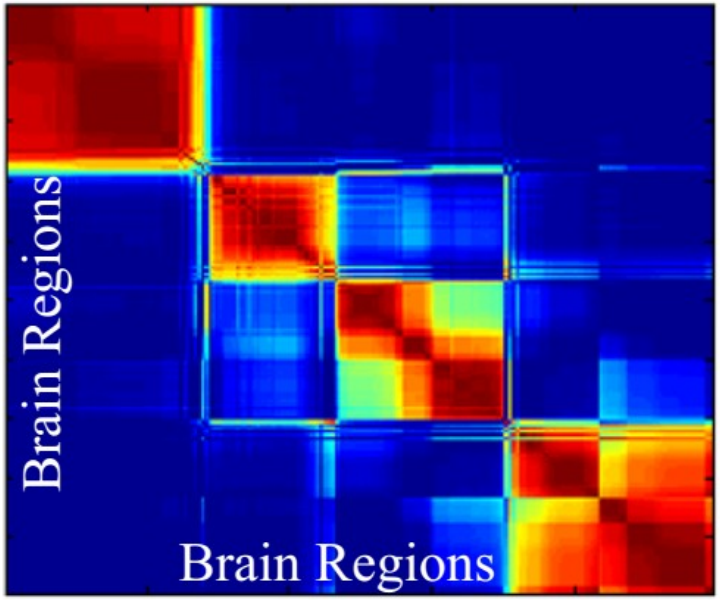


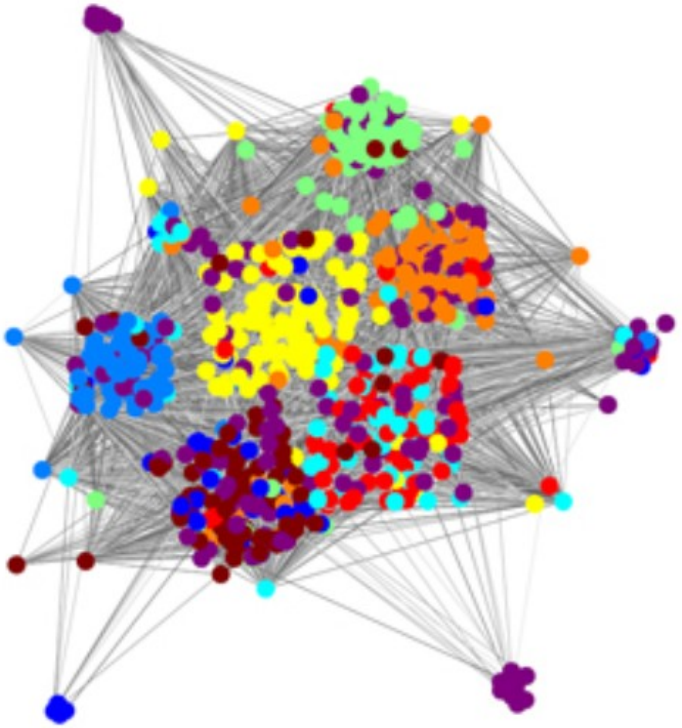
Neuronal subnetworks

Hierarchy and modularity is observed at large scale: several subnetworks are responsible for arousal, attention, positive/negative valence, perception.

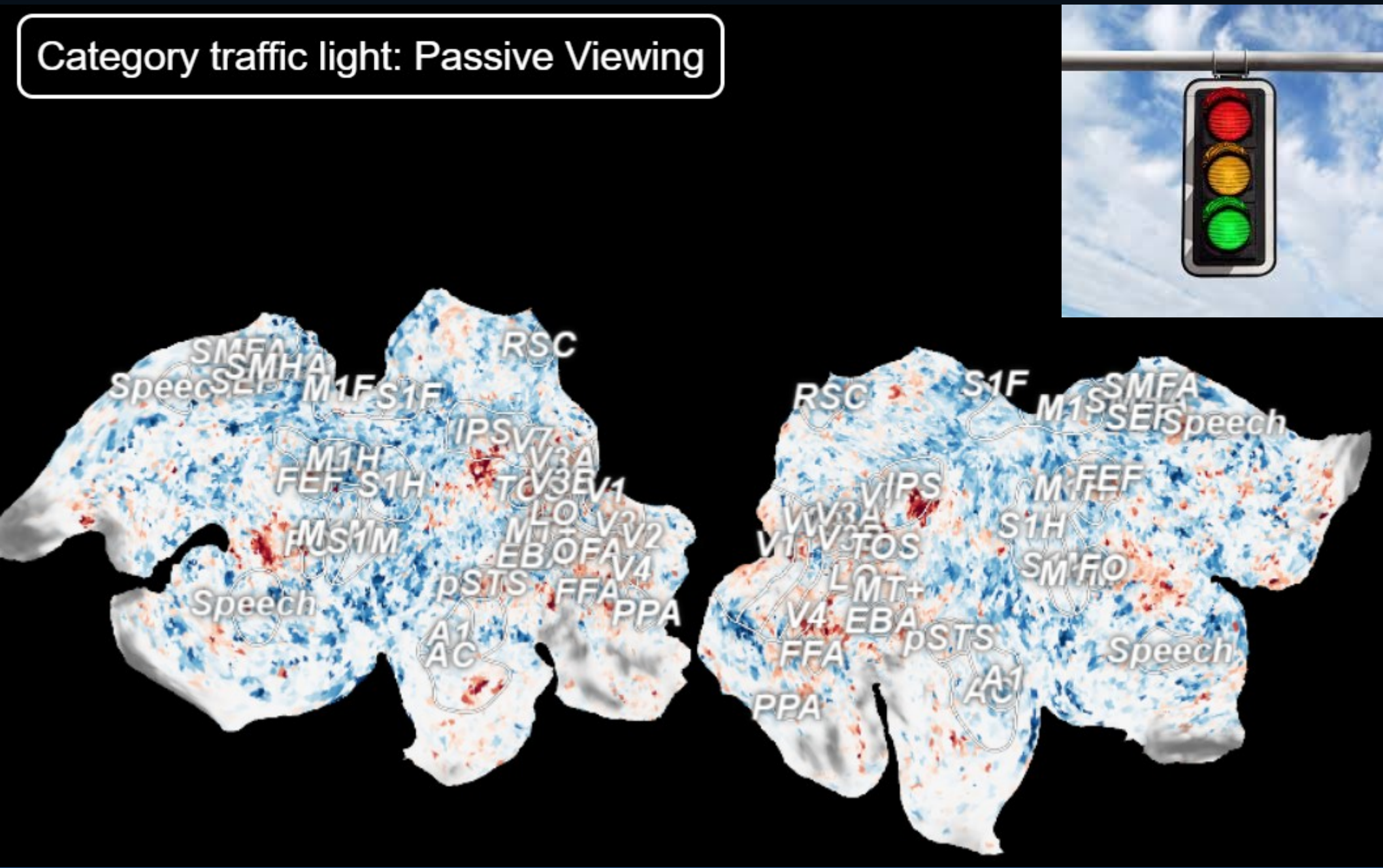
At the microcircuit level similar hierarchy and modularity is seen.

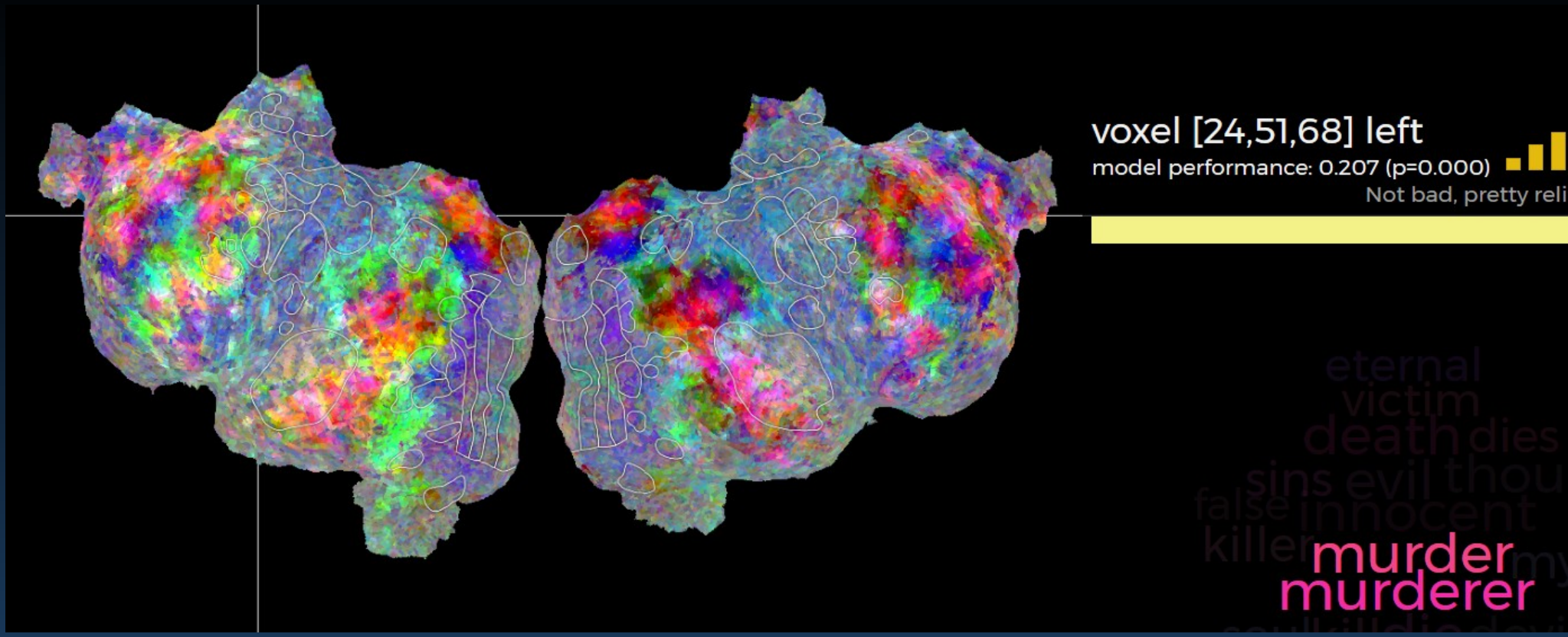






Interpretation for simple objects is easy: IPS – visual attention, V4 – color, AC – object recogniton.



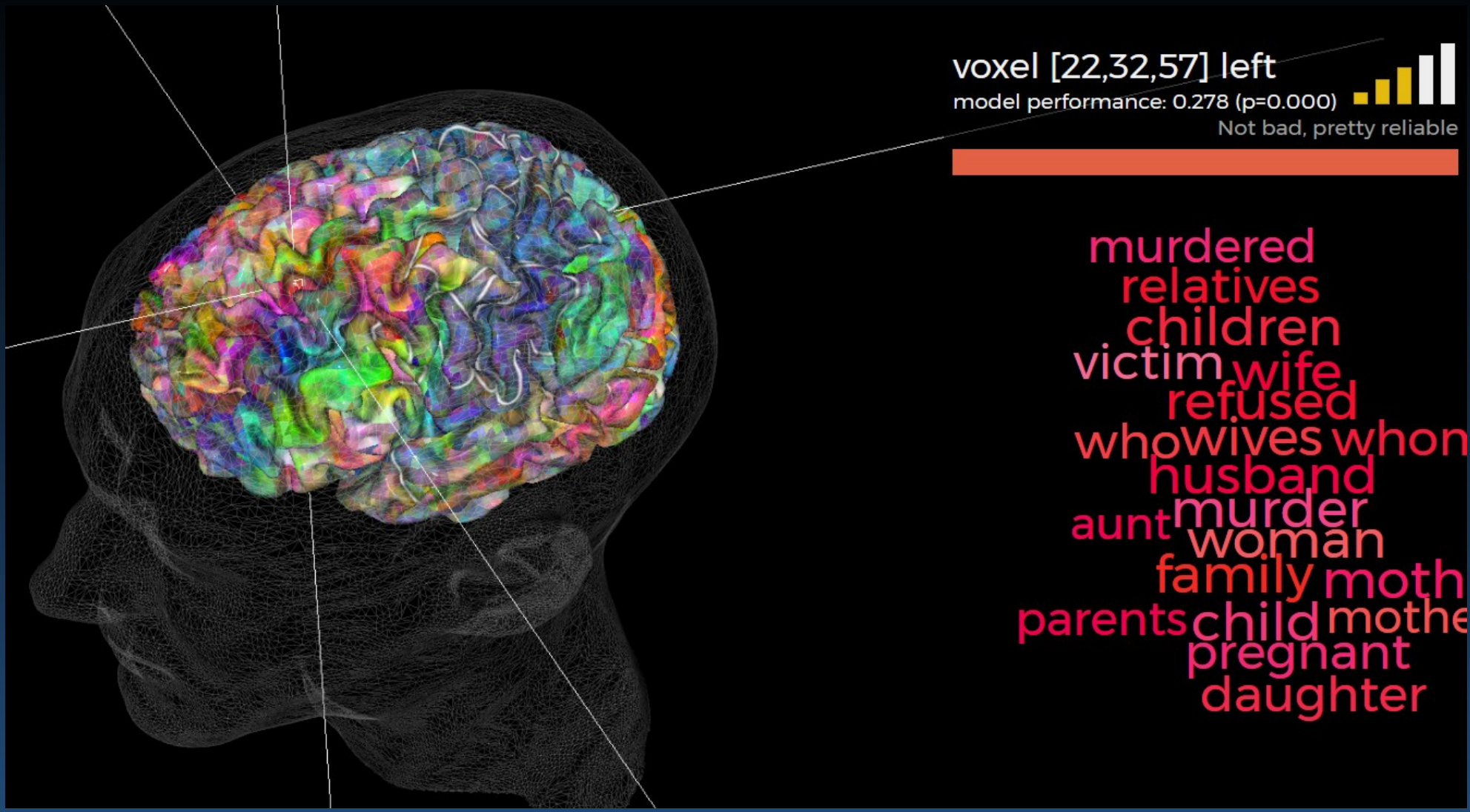


Whole fMRI activity map for the word “murder” shown on the flattened cortex.

Each word activates a whole map of activity in the brain, depending on sensory features, motor actions and affective components associated with this word.
Why such activity patterns arise? Brain subnetworks connect active areas.

http://gallantlab.org/huth2016/ and short movie intro (A. Huth, Nature).

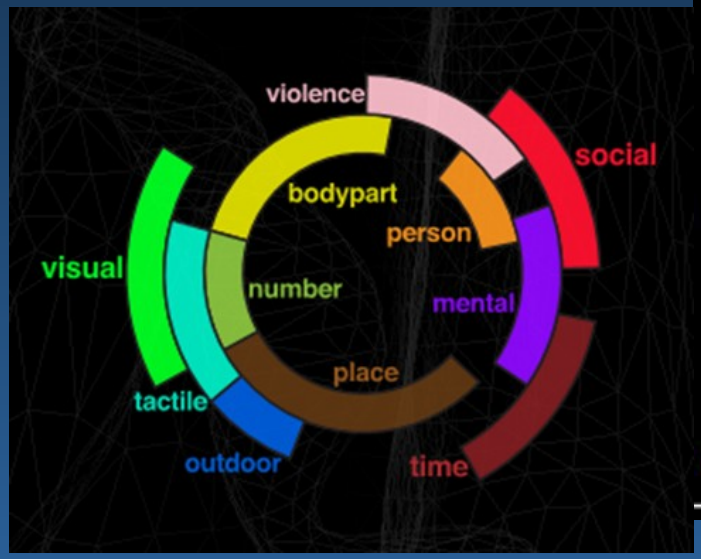
Can one do something like that with EEG or MEG?

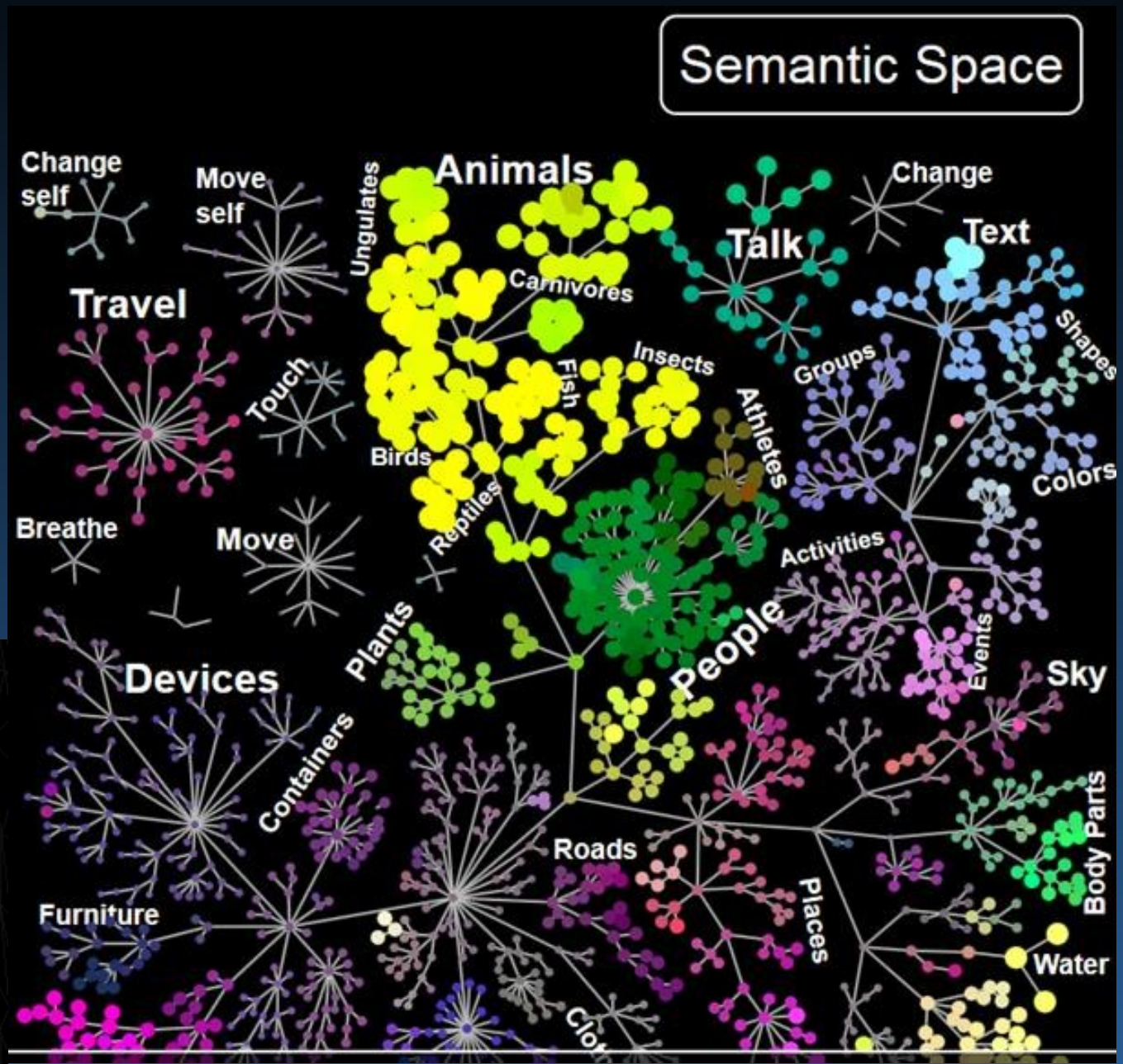


Each voxel responds usually to many related words, whole categories.
http://gallantlab.org/huth2016/
Huth et al. (2016). Decoding the Semantic Content of Natural Movies from Human Brain Activity. Frontiers in Systems Neuroscience 10, pp. 81

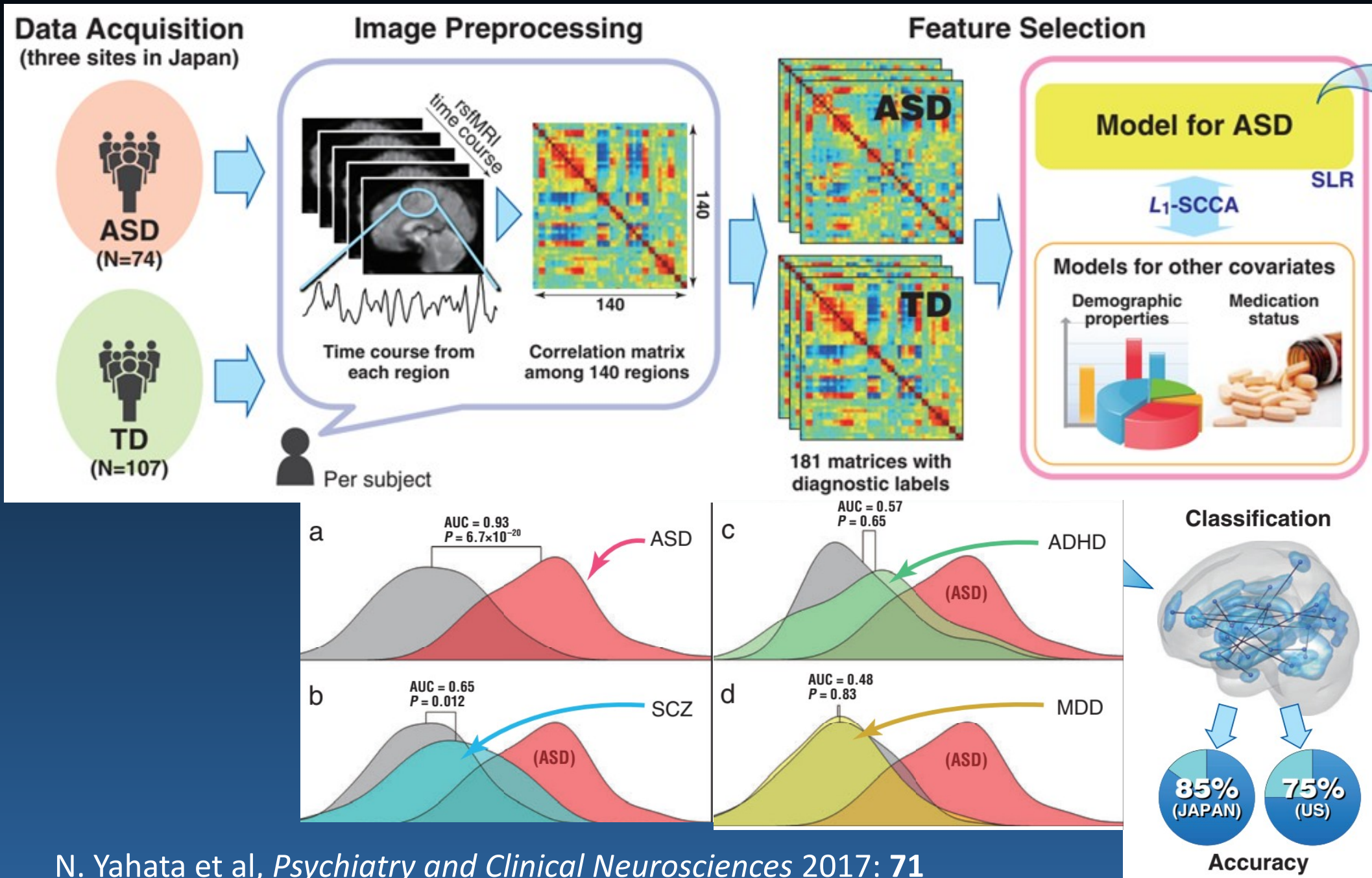
Semantic neuronal space

1700 words in the semantic space are grouped by similarity. Words activate specific ROIs, similar words create similar maps of brain activity. Video or audio stimuli, fMRI 60.000 voxel). Gallantlab,Berkeley.





Biomarkers from neuroimaging



N. Yahata et al, *Psychiatry and Clinical Neurosciences* 2017: **71**

Nicole Speer et al.
Reading Stories Activates Neural Representations of Visual and Motor Experiences. *Psychological Science* 2009

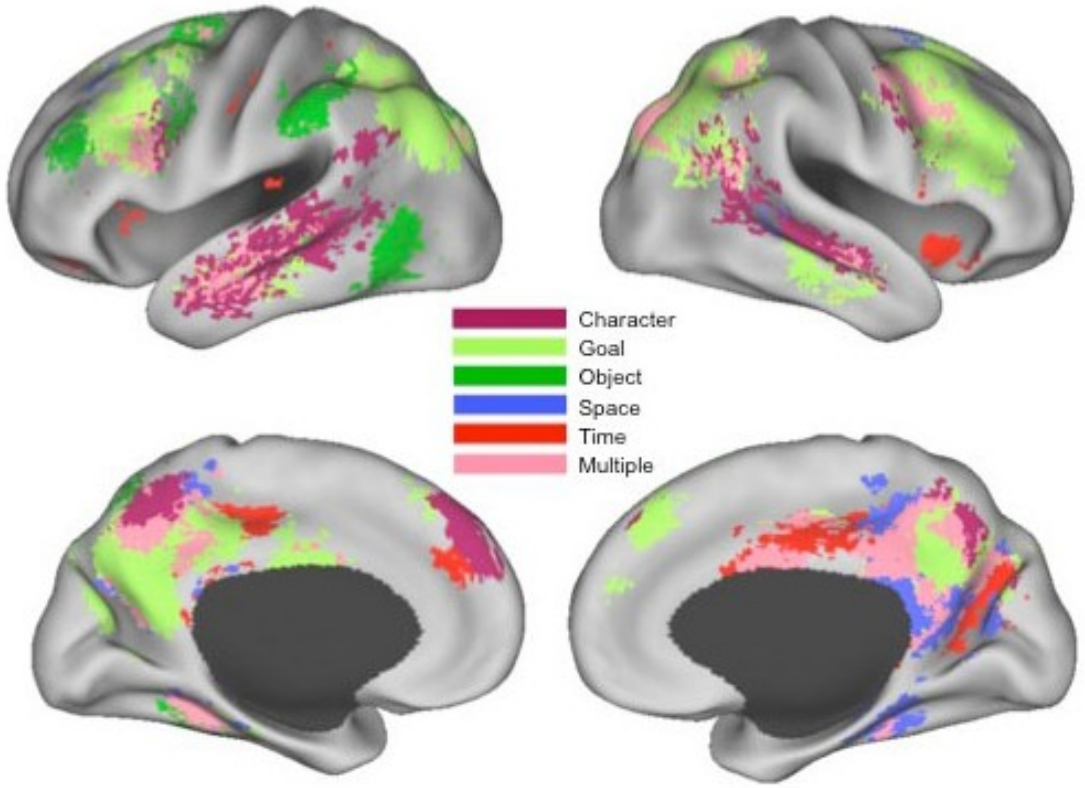
Automatic segmentation of experience is the basis of perception, facilitates planning, memory, association of information. Transitions between segments result from important observations in the current episode, entering new objects, places, goals, interactions, like in a movie.

Coherent narratives are recalled and coordinated by hippocampus (Cohn-Sheehy et al., Curr. Bio. 2021).

A

Clause	Cause	Character	Goal	Object	Space	Time
...[Mrs. Birch] went through the front door into the kitchen.	●				●	
Mr. Birch came in	●	●			●	
and, after a friendly greeting,	●					●
chatted with her for a minute or so.	●					●
Mrs. Birch needed to awaken Raymond.		●				
Mrs. Birch stepped into Raymond's bedroom,			●		●	
pulled a light cord hanging from the center of the room,				●		
and turned to the bed.						
Mrs. Birch said with pleasant casualness,						
"Raymond, wake up."						
With a little more urgency in her voice she spoke again:						
Son, are you going to school today?						
Raymond didn't respond immediately.		●				●
He screwed up his face			●			
And whimpered a little.						

B



Hippocampus
and formation of memories

Hippocampus

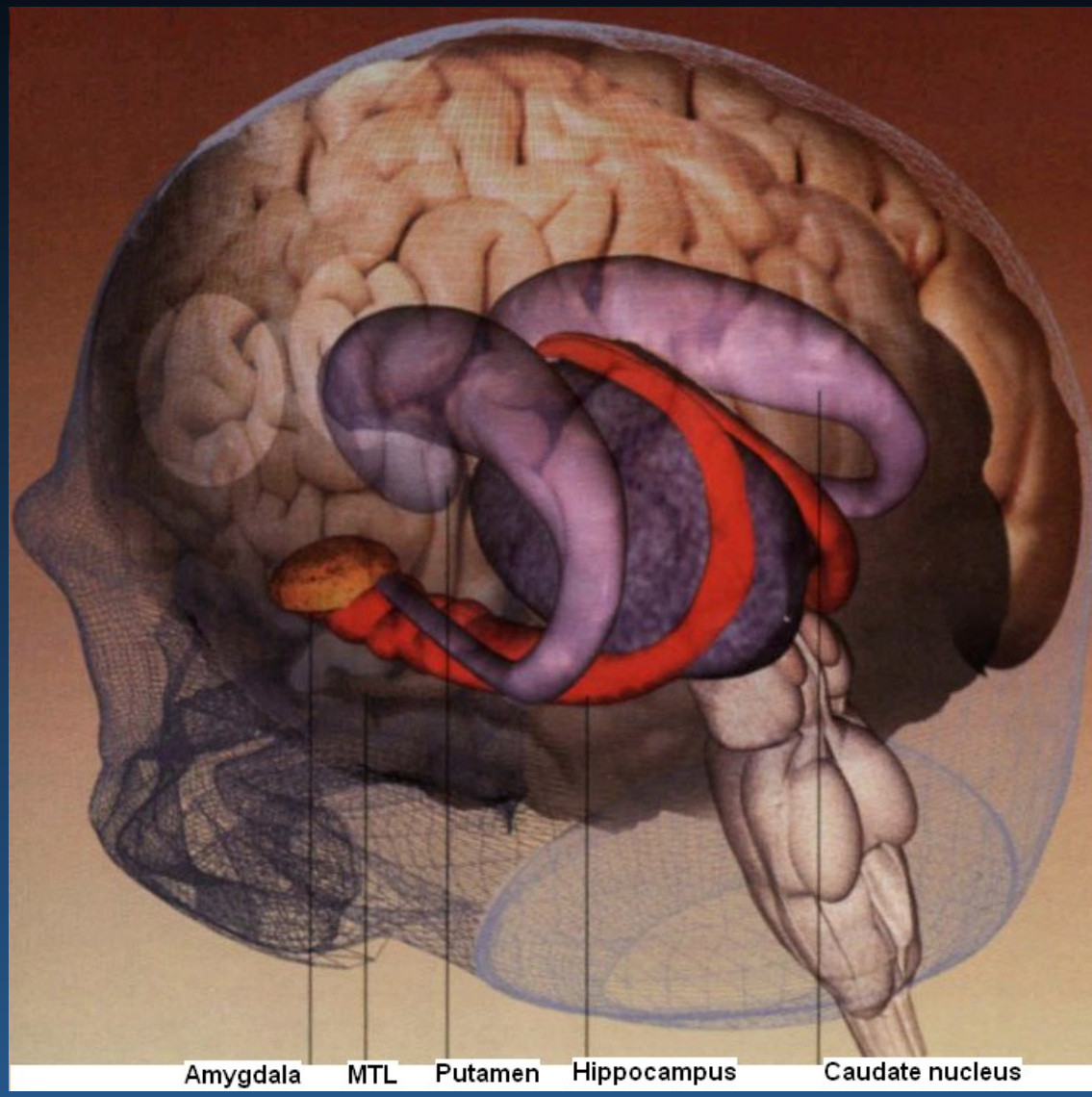
Hippocampal formation (red color) involves entorhinal, perirhinal and parahippocampal cortex.

Signals reach many uni-modal and multi-modal association cortical areas.

Hippocampus has many functions:

- Spatial memory – grid, head direction, border, speed cells => place cells.
- Consolidation of long-term memory
- Synaptic learning: Long-Term Potentiation.

Animation of memory formation



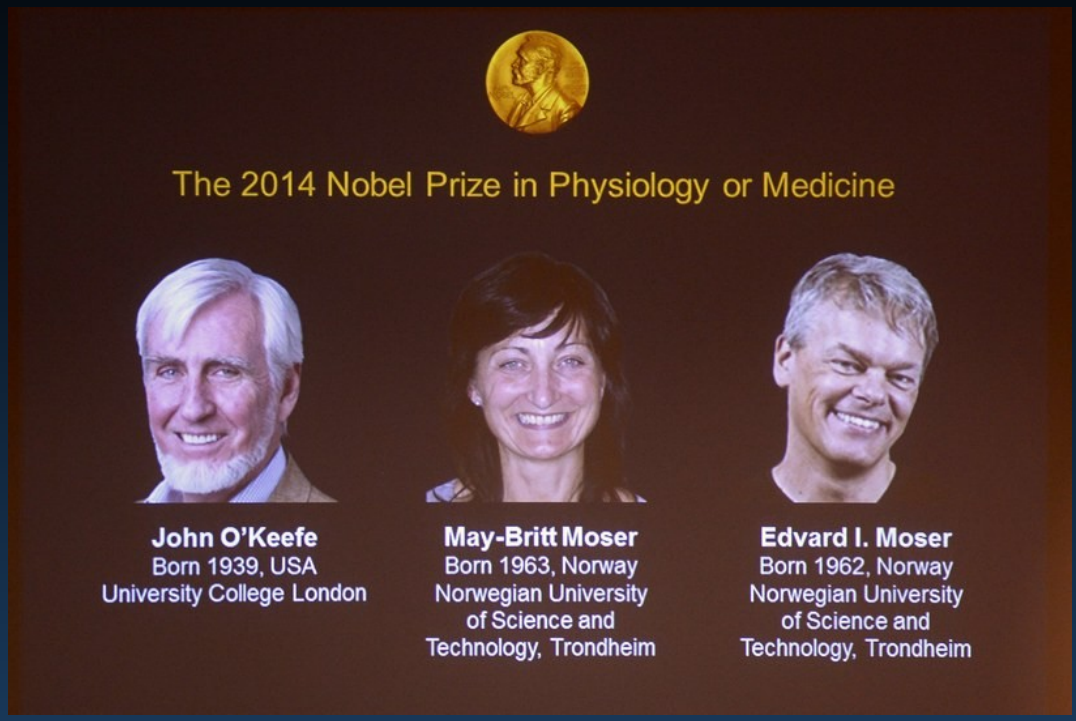
Spatial memory

Nobel Prize Lectures:

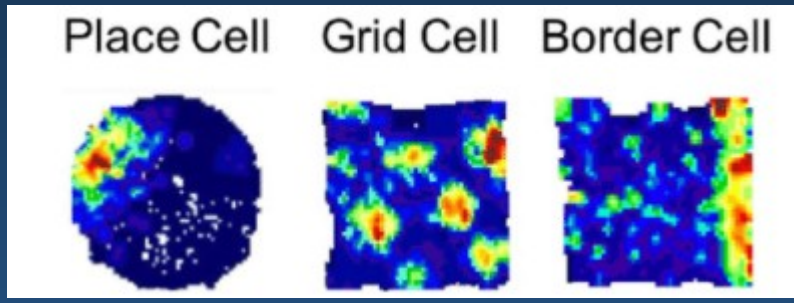
John O'Keefe. The Hippocampus as a Cognitive Map.

May-Britt Moser. Grid Cells, Place Cells and Memory.

Edvard Moser. Grid Cells and the Enthorinal Map of Space.

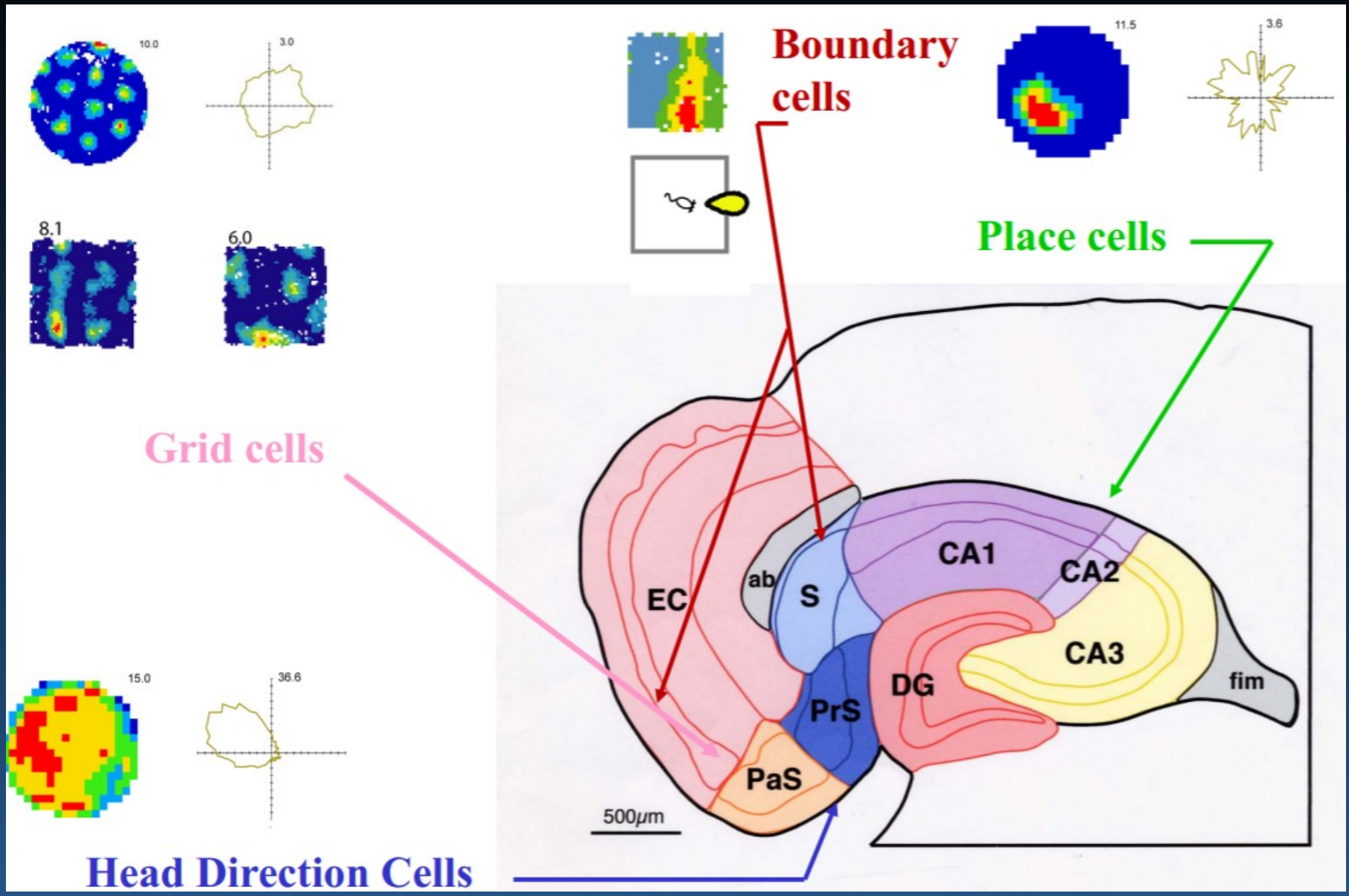


Brain GPS system: grid cells, the head direction cells, and the border cells, all project to hippocampal place cells. The mapping system provides two independent strategies for locating places, one based on environmental landmarks, and the other on a path integration system which uses information about distances traveled in particular directions.



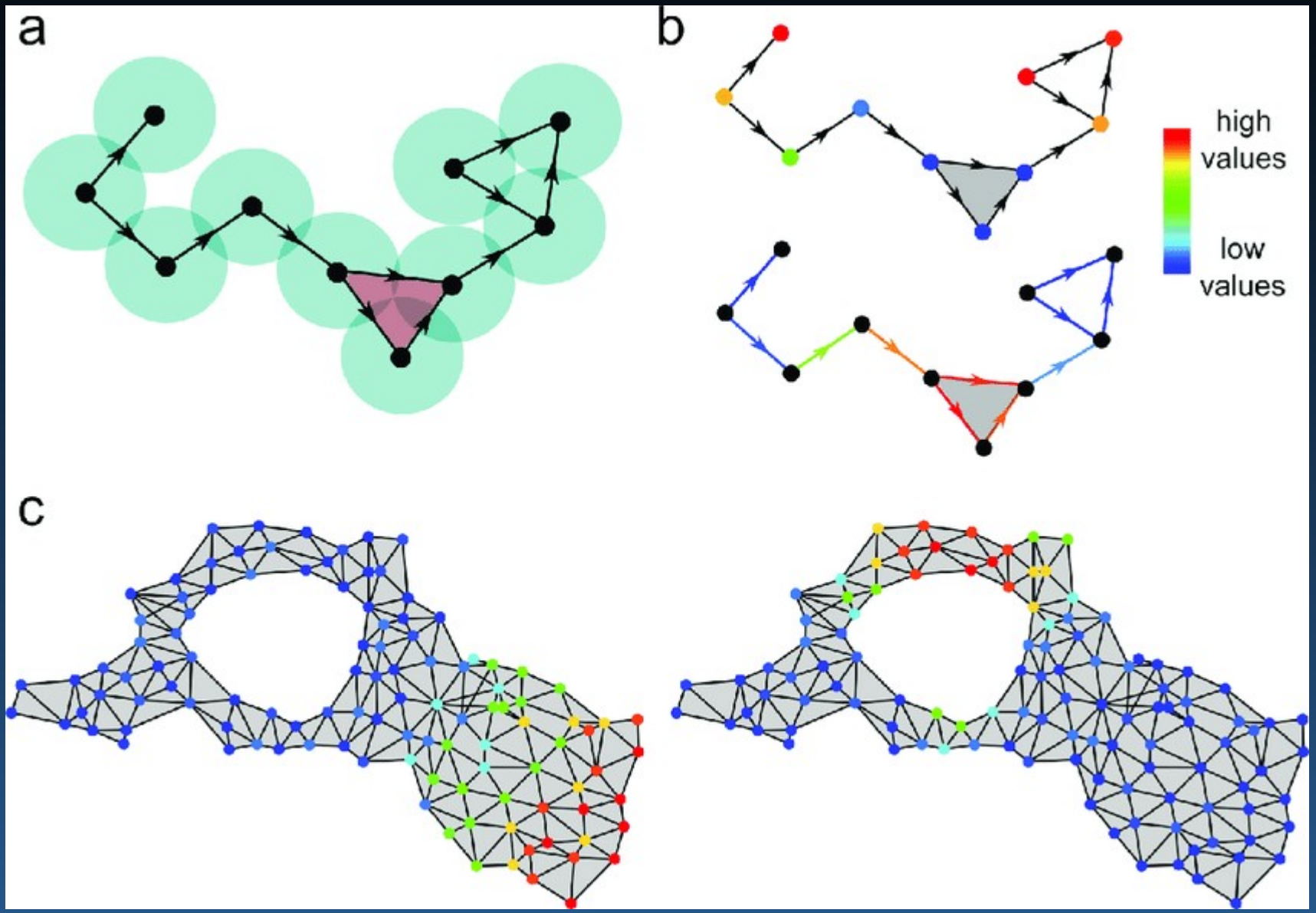
Path integration:
Ants measure distance counting steps. Rats use "speedometer".

Cells in hippocampal formation

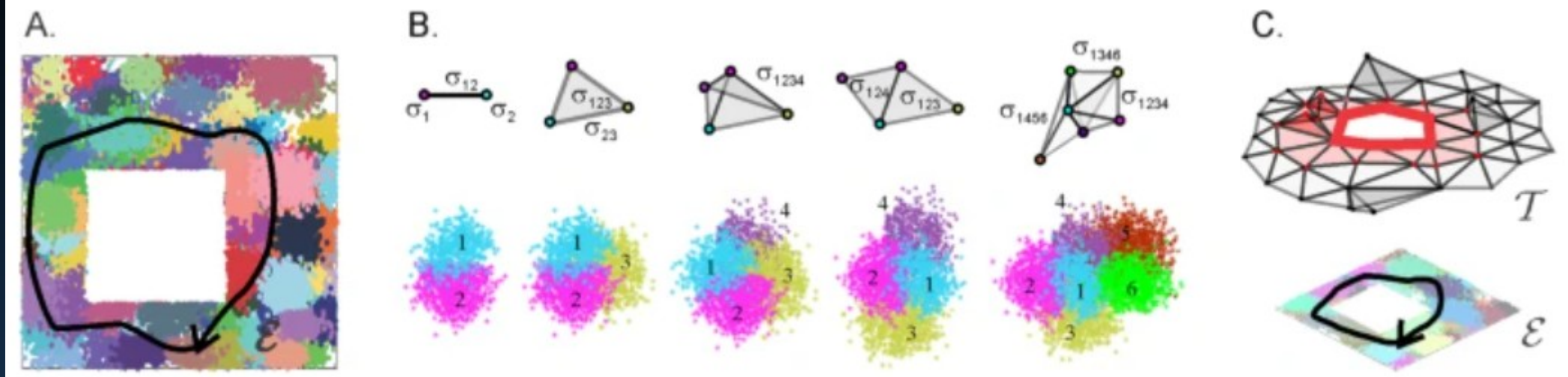


Spatial memory: different cells are active in different places, one based on environmental landmarks, the other on a path integration of distances traveled in particular directions.

Formation of place cell complexes

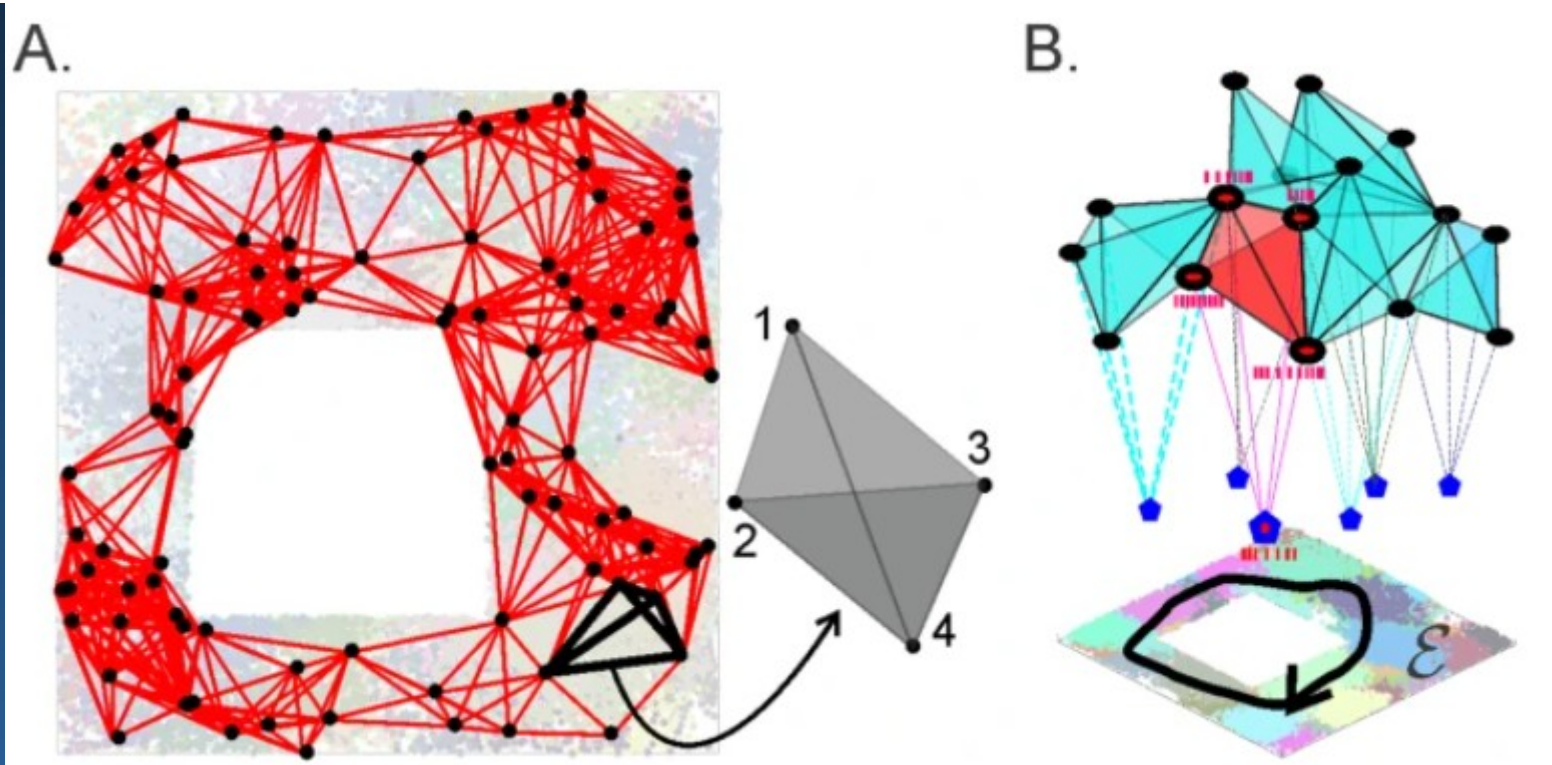


Simplicial complexes in hippocampus



Hippocampal place cell ensembles produce a cognitive map of the environment. Weak synaptic connections produce topological defects in the large-scale representation of space.

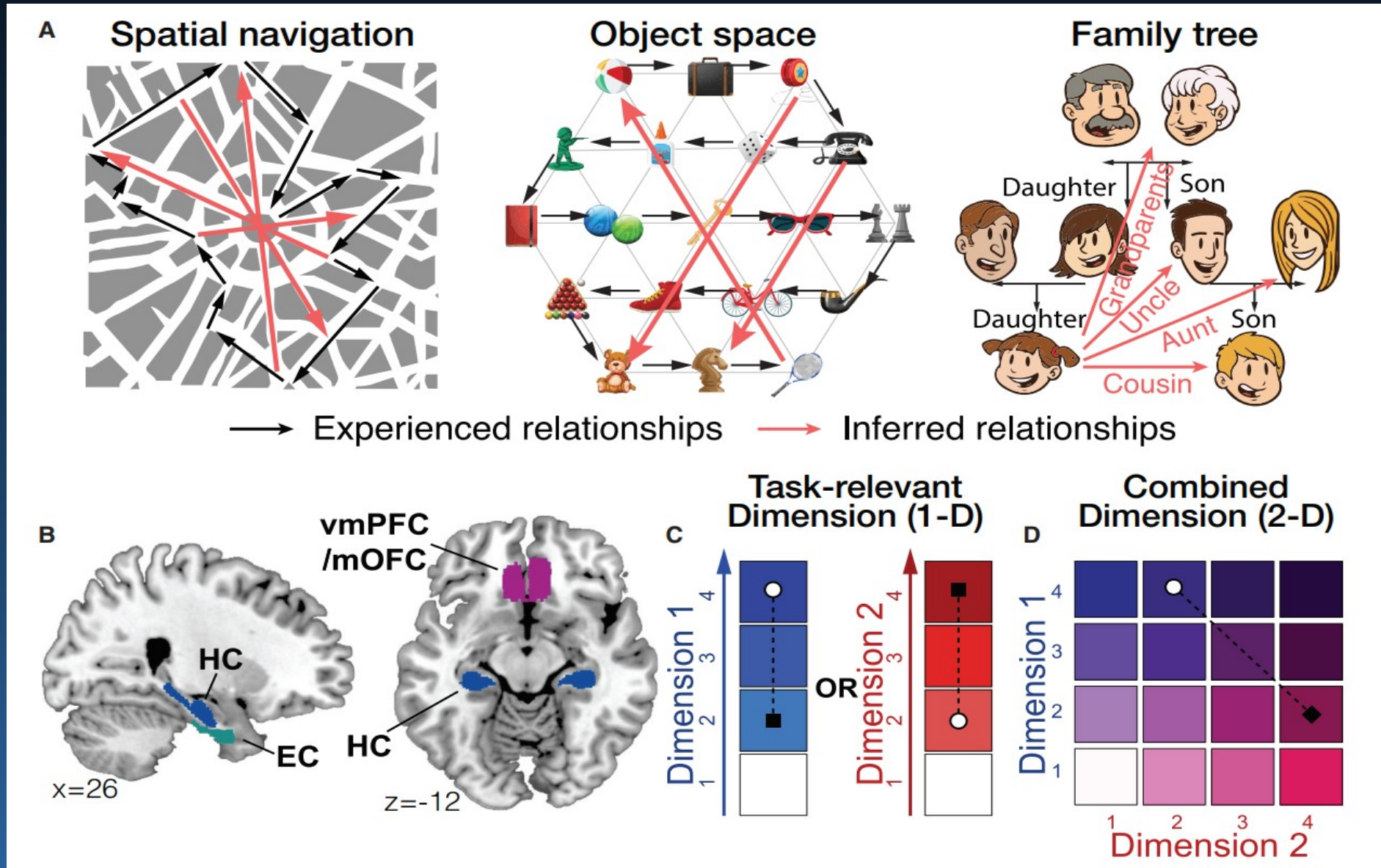
Dabaghian, Y. (2019, 2021). From topological analyses to functional modeling. Sci. Rep.



Abstract Cognitive Maps

S.A. Park et al, Map Making: Constructing, Combining, and Inferring on Abstract Cognitive Maps. Neuron 2020.

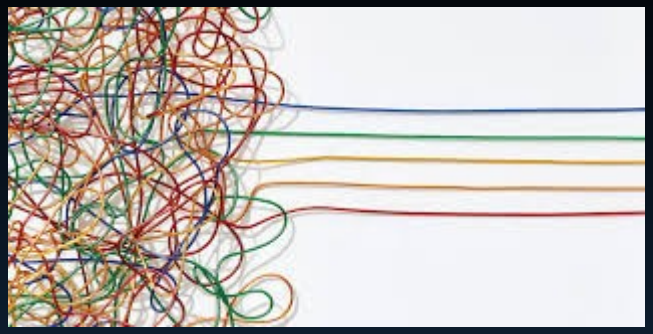
Black lines – visited, red = inferred. Places and relations => object relations.



From May-Britt Moser Nobel lecture

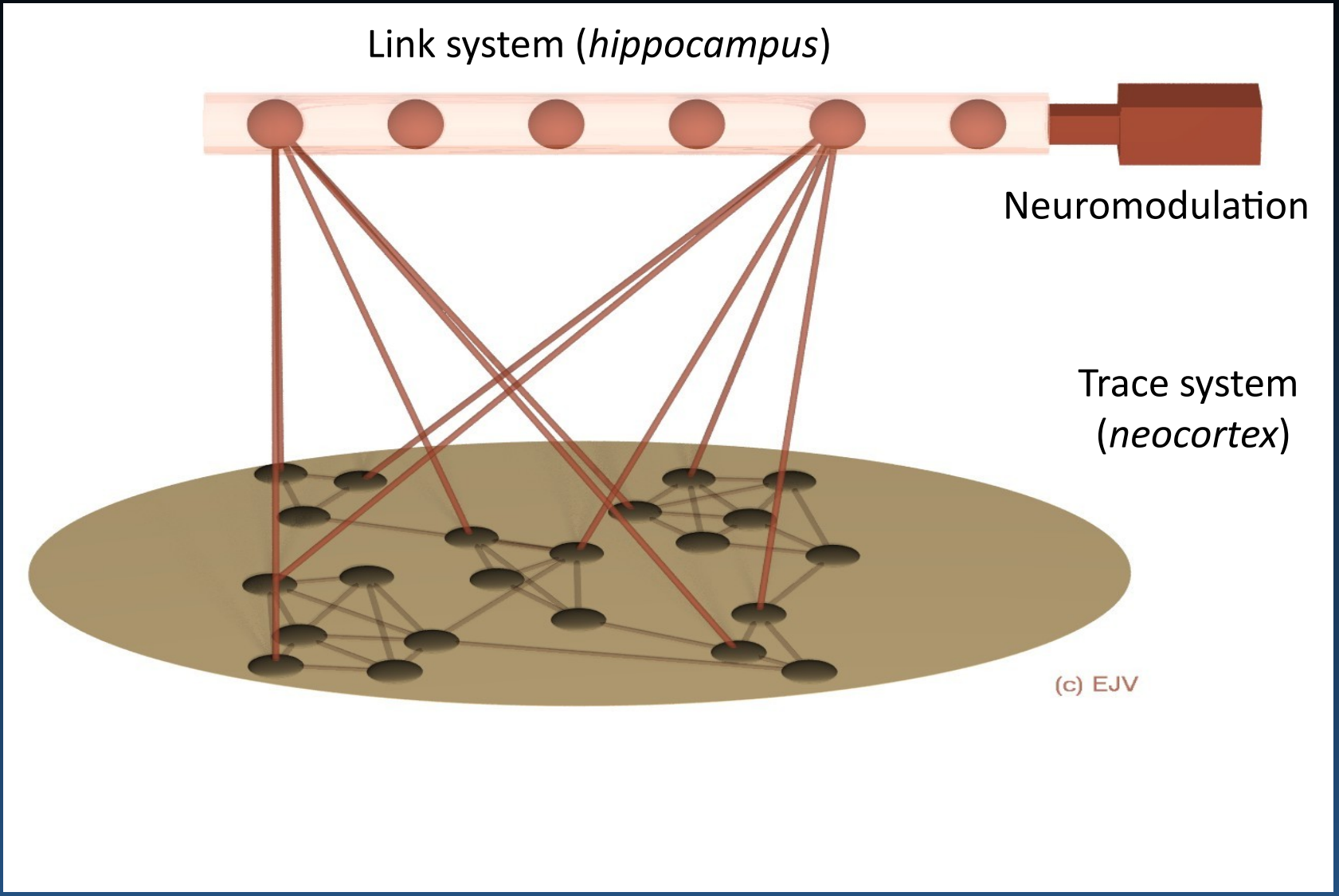
Space is used as a framework for storing memories

Mnemonic techniques, memory palace



We need to simplify ...

Trace-link model (J. Murre)



Hippocampus may recreate activation of the cortex => episodic memory.
Brain stem and emotion-related areas (nucleus accumbens, amygdala) provide increase plasticity. This simple model explains many features of amnesia.

Computational Models

Biologically plausible models of neural networks should use neurons with at least 3 types of ion channels.

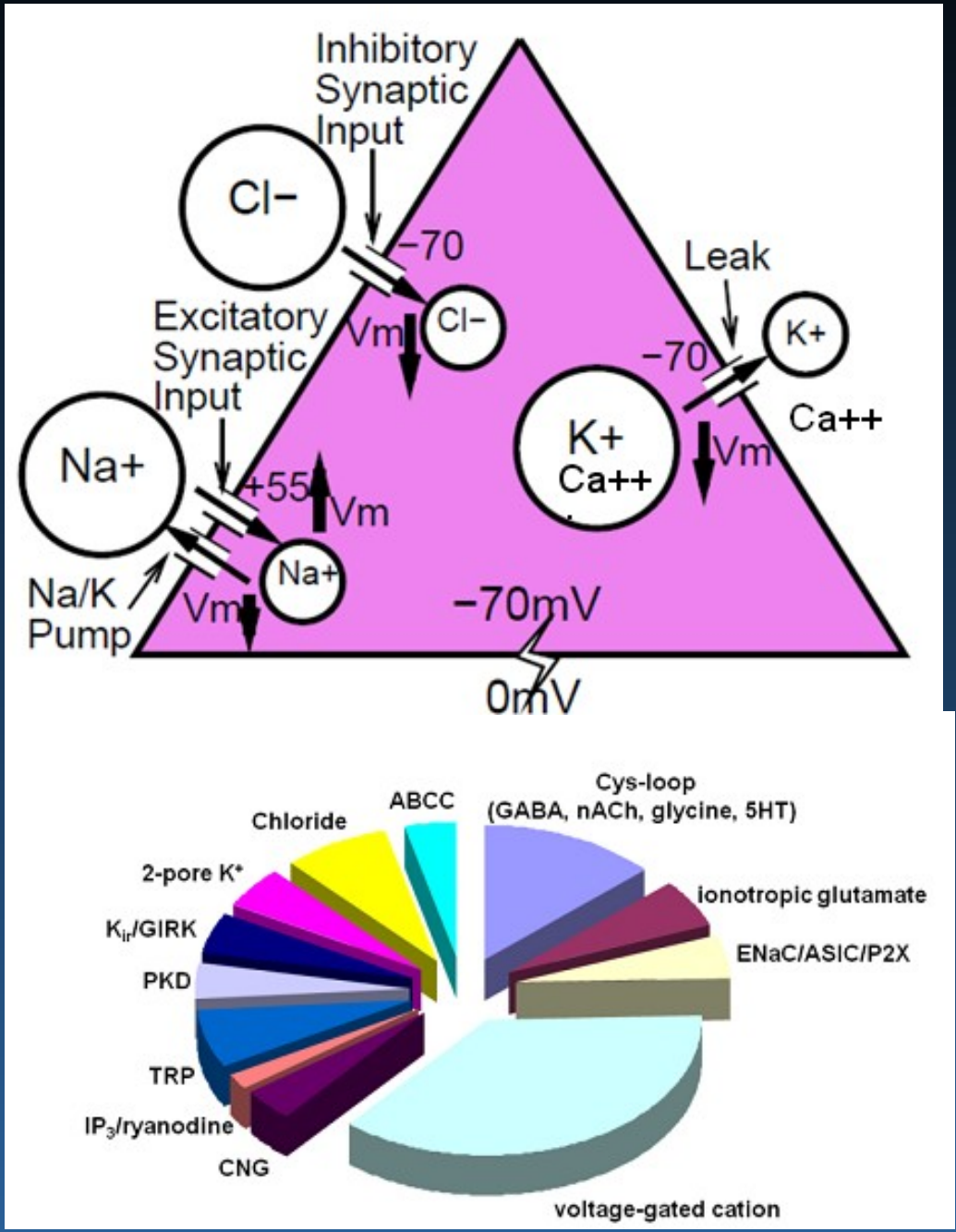
100 mln ions flow through each channel/sek. Neurons has 10 000 channels.

Intracellular calcium builds up as a function of activation.

Initial focus is on the leak channels, controlling time spend in basins of attractors (trapping time).

Ex. 2-pore K^+, proteins and genes may be identified and their dysfunctions analyzed in various disease.

Models of word associations: sequence of spontaneous thoughts.

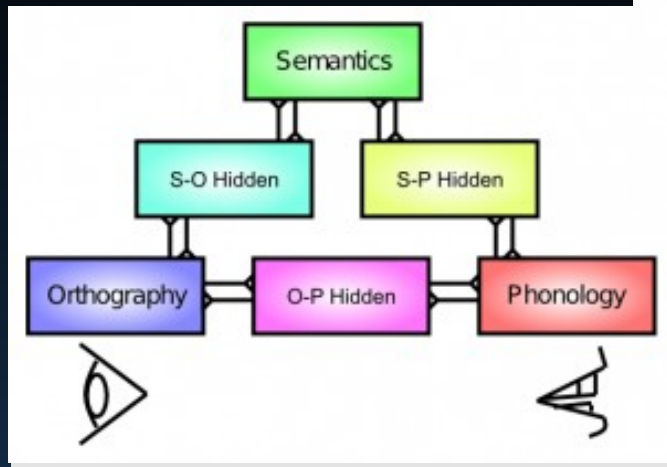


Model of reading & dyslexia

3-way or triangle model of reading:

orthography – visual perception,
phonology – spoken (motor) output,
semantics – meaning as distributed activity over 140 microfeatures.

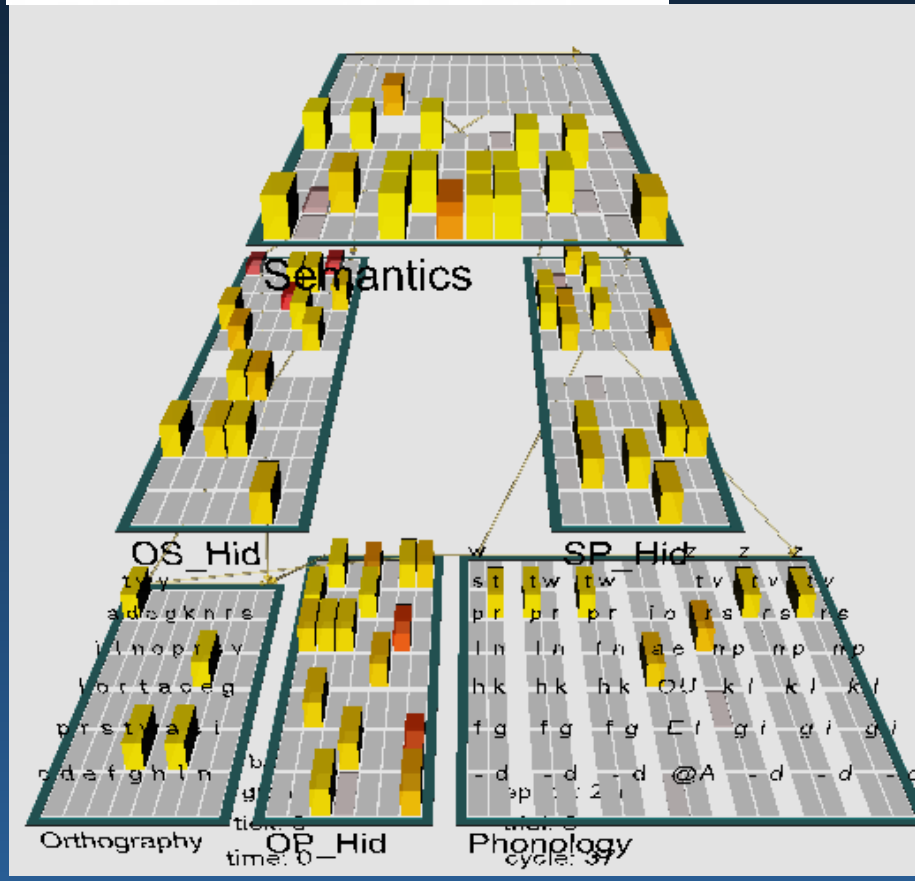
In between are transformation layers.

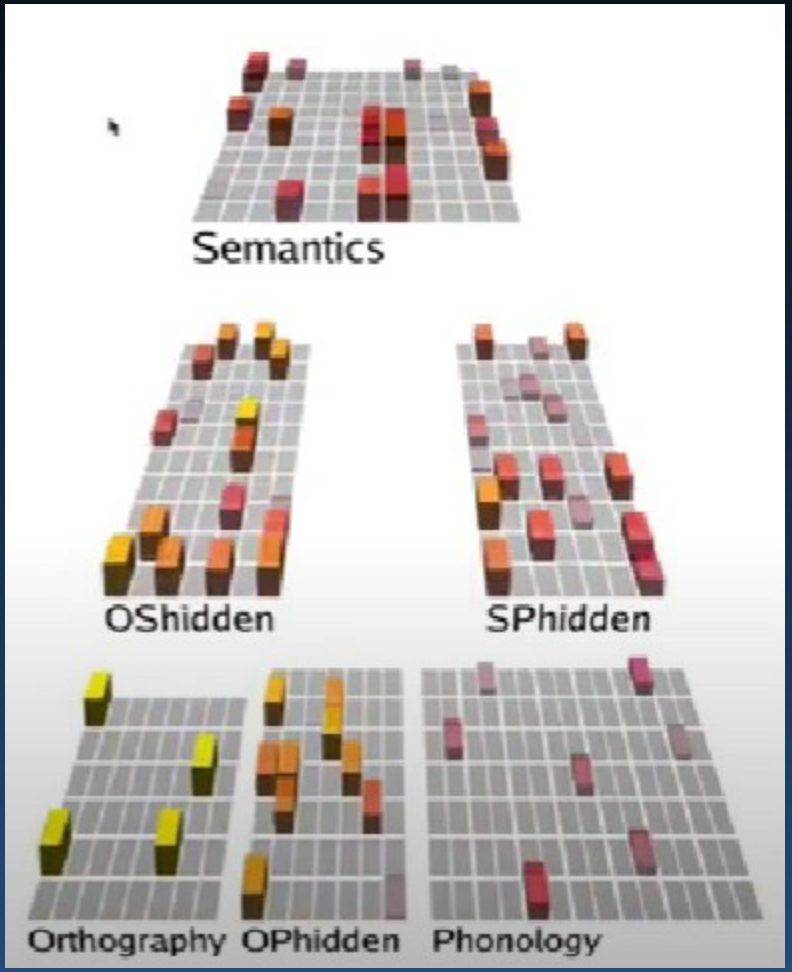


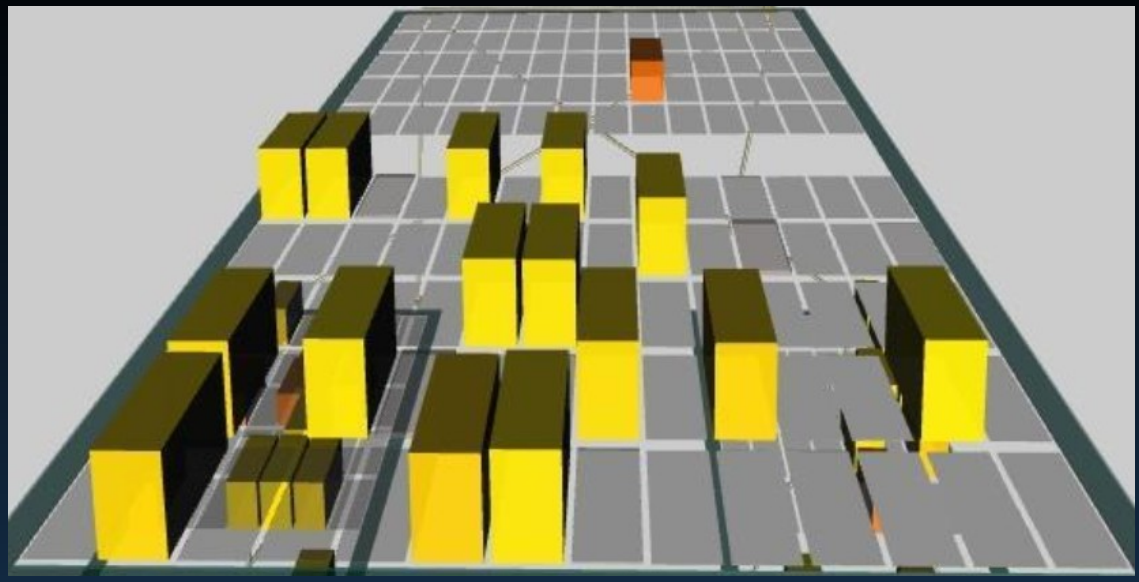
Visual input (orthography) => speech output directly via projections to phonology or indirectly, with comprehension, via orthography => semantics => phonology.

Representations of word meanings are distributed across most of the cortex. Damage to different pathways can account for properties of acquired dyslexia.

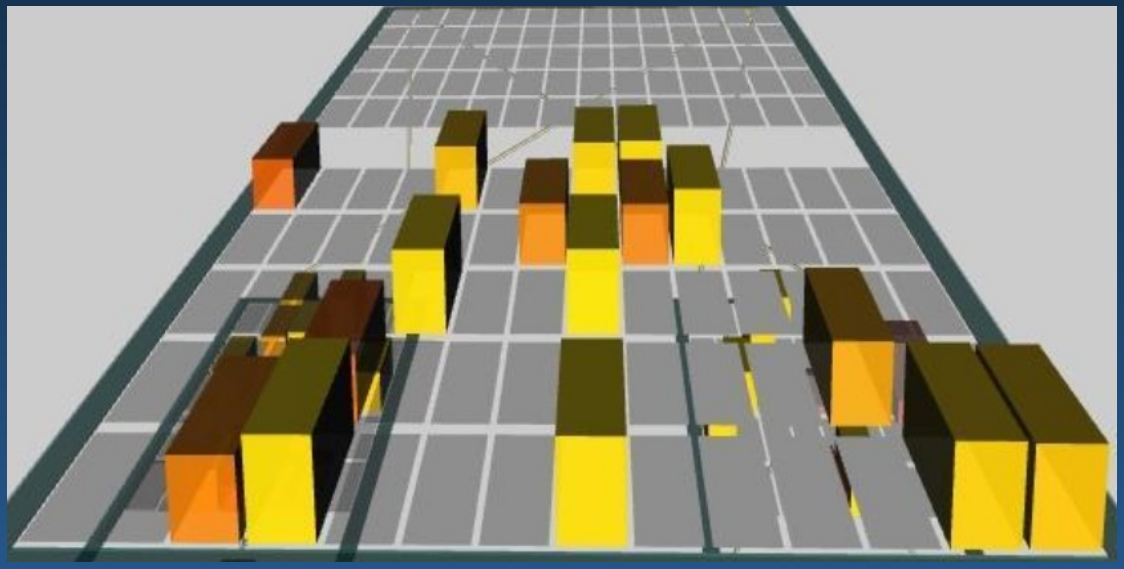
Learning: mapping one of the 3 layers to the other two.



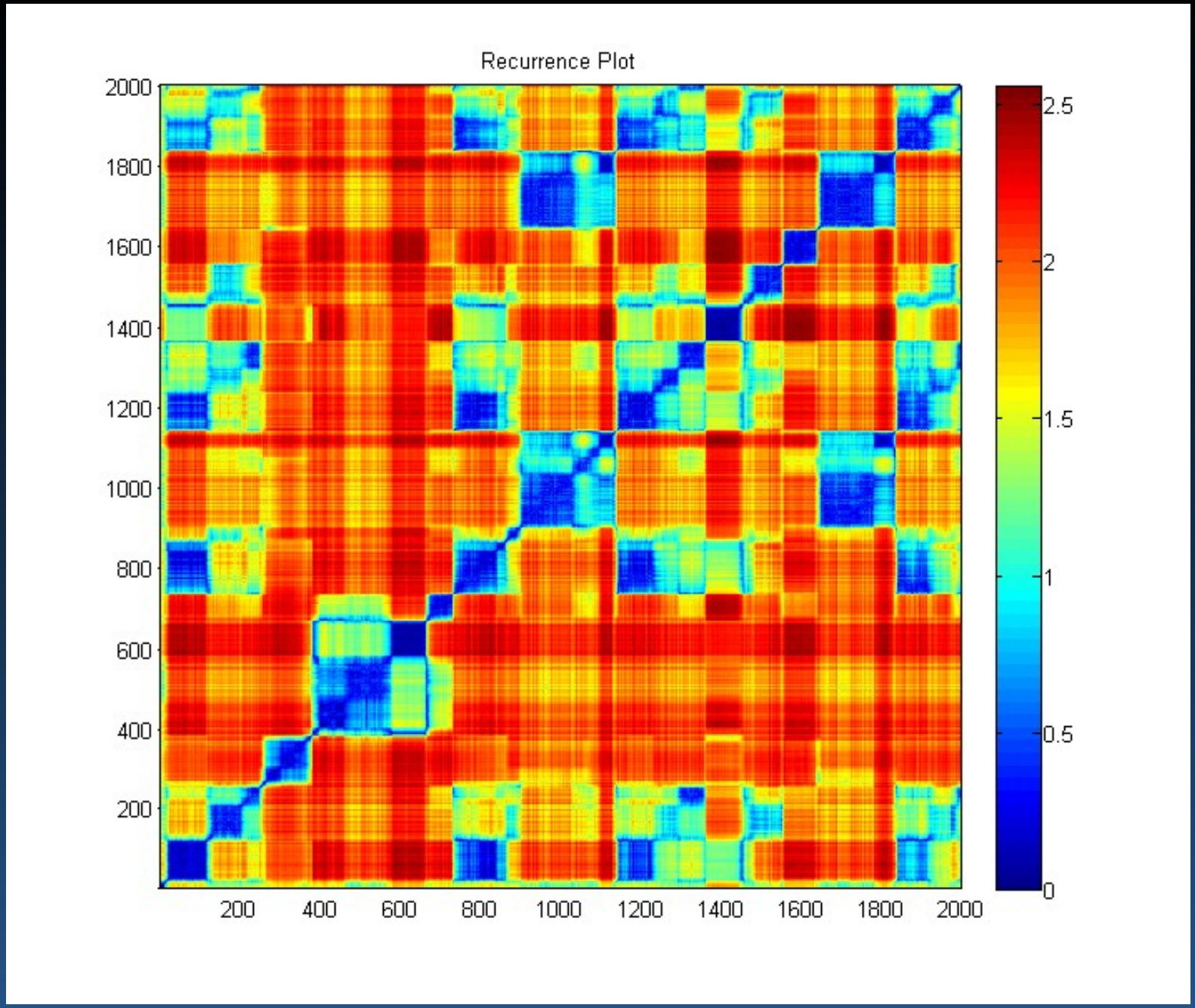




Transition from “case” to “rope”.

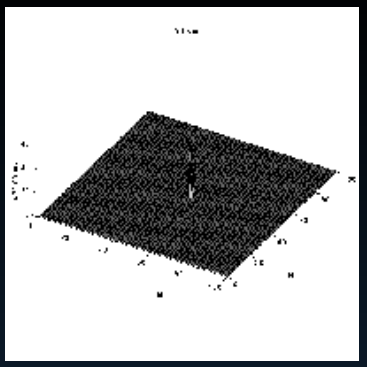


Emergent neural simulator, at compcogneuro



Recurrence plots show color map of distance between ||X(t)-X(t')||. If the activation pattern is quasi-stable distance is low (dark blue squares along diagonal). After 100 steps starting from the word „gain" system makes transition to another attractor state.

Basins of attractors

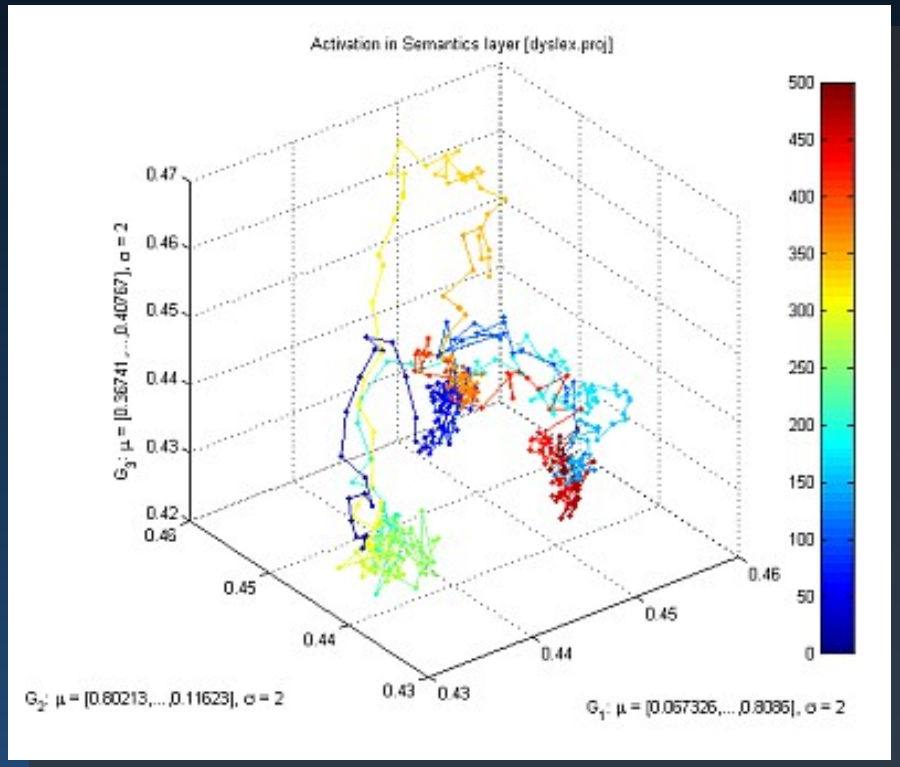


Groups of highly active neurons synchronize, their activation fluctuates around creating specific distribution maps, inhibiting competing groups of neurons.

Normal case: relatively large attractor basins, short dwell times, fast transitions from one basin of attraction to another, creating “stream of distinguishable internal states”.

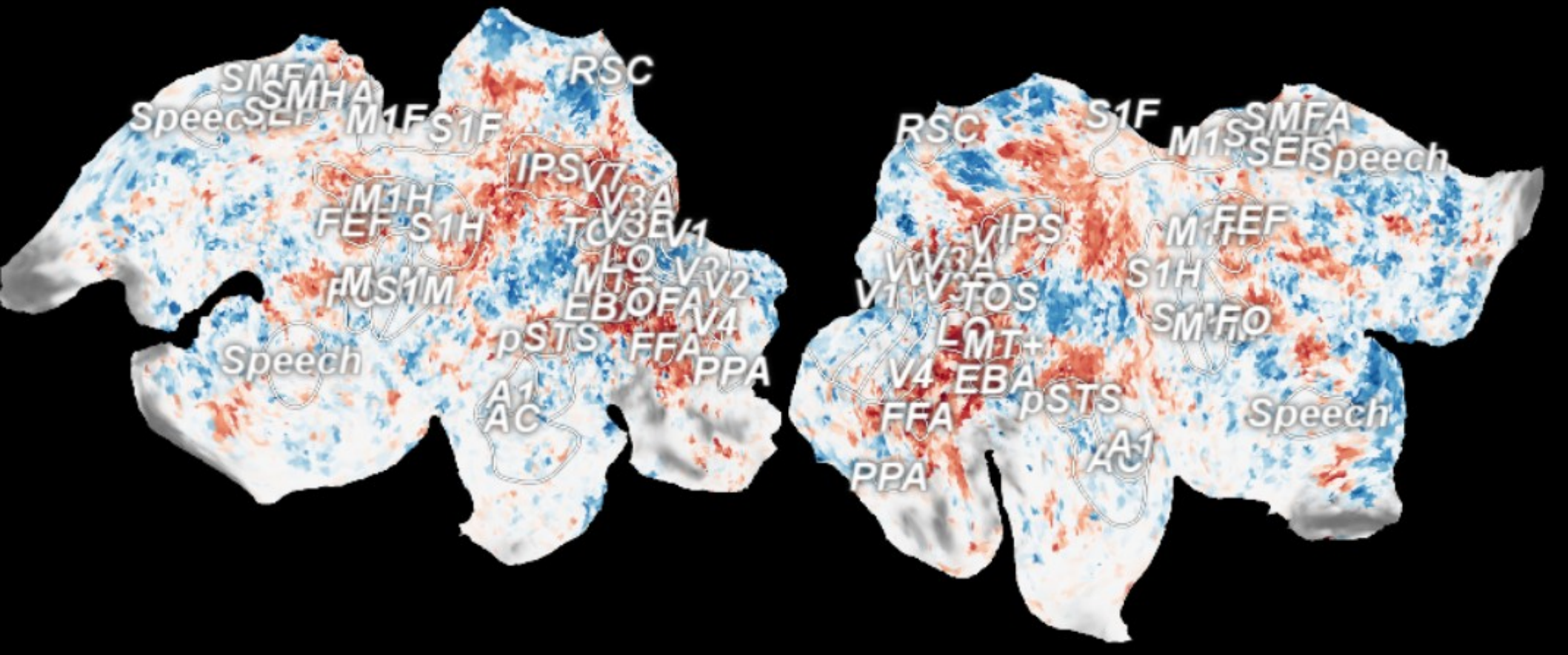
Each point here ⇔ vector in 140 dim.

Brain has about 3 mln minicolumns in the cortex, corresponding to units in computational model, participating in creating quasi-stable patterns.

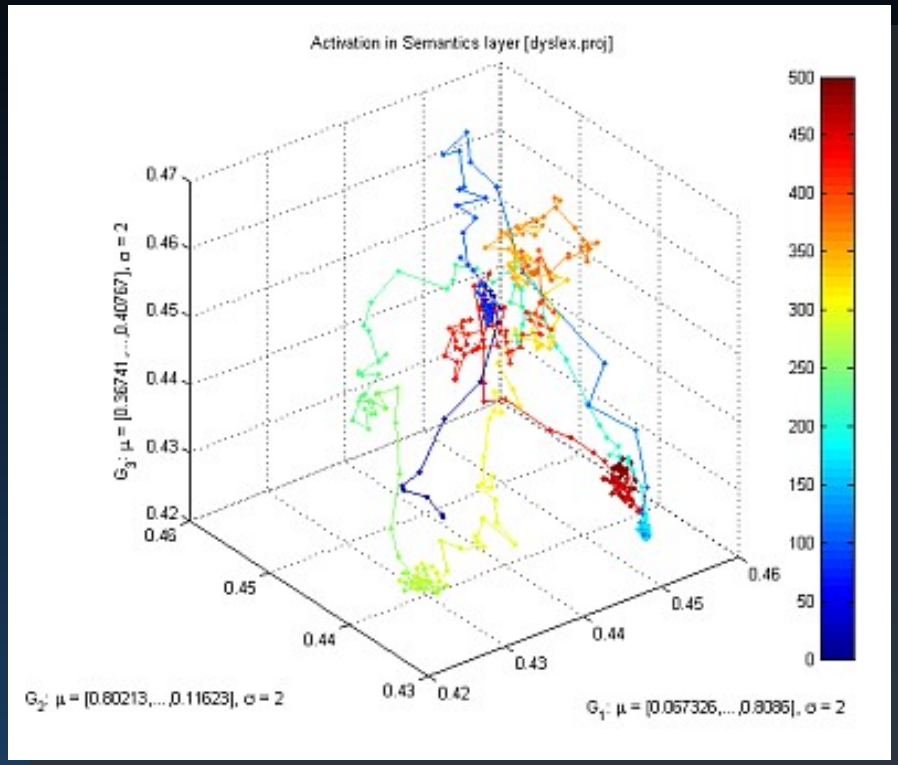


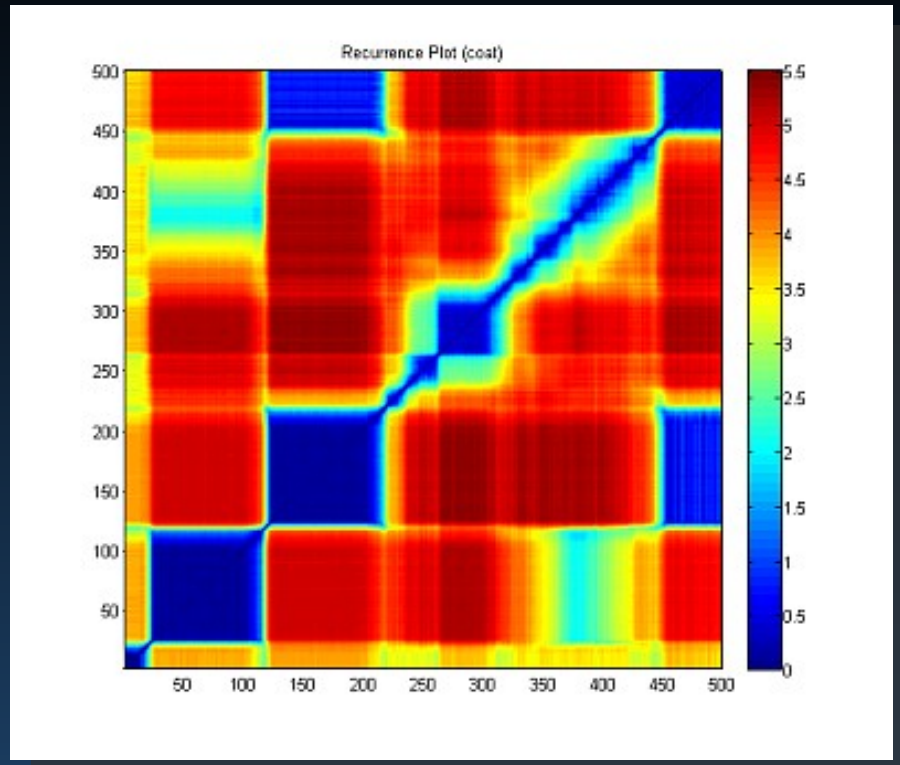
Accessible basins of attractors = available mental states that can be categorized and identified. They shrink and vanish as neurons desynchronize due to the fatigue; this allows other neurons to synchronize, leading to new mental states (thoughts).

Category zebra: Passive Viewing



Fast transitions



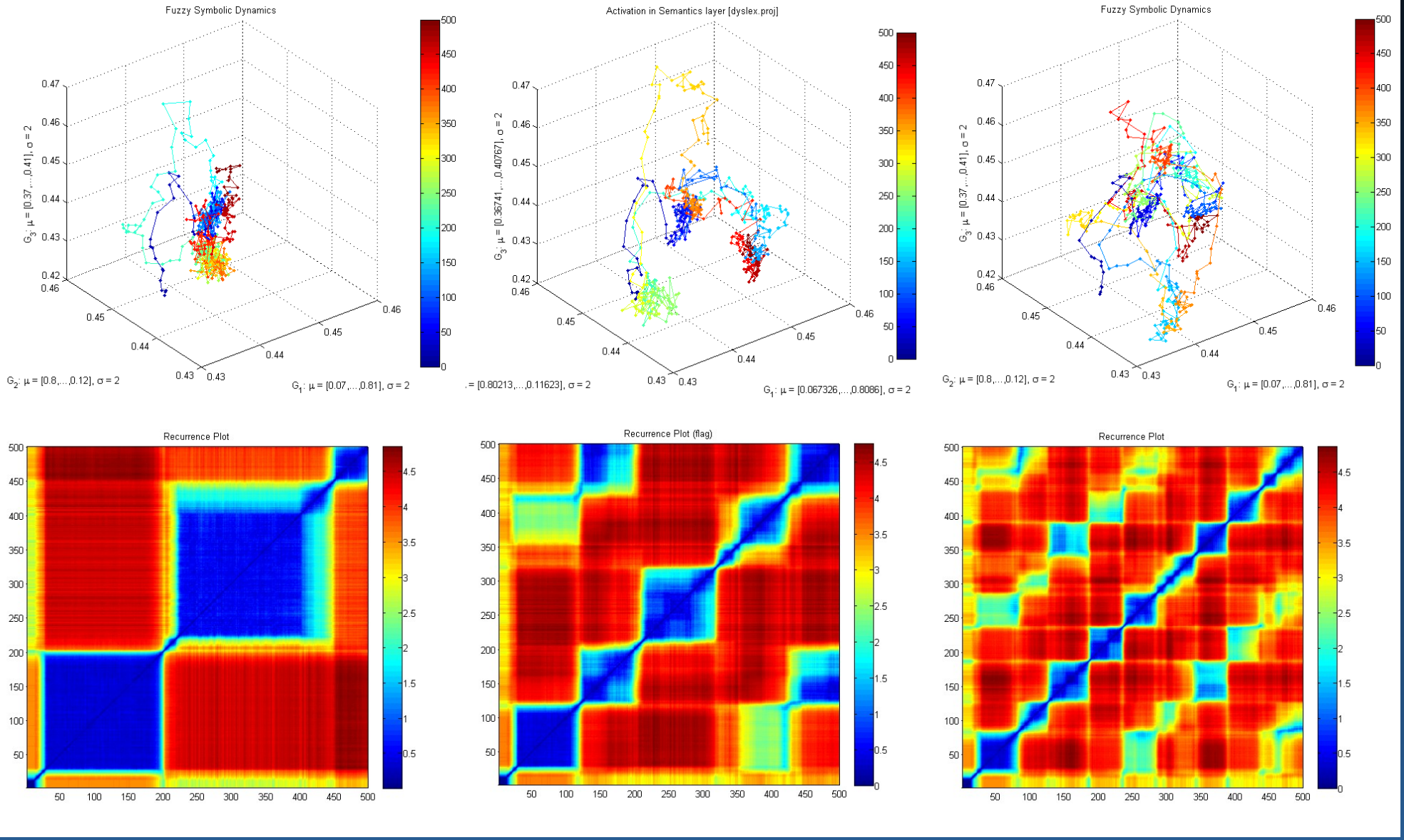


Strong leak currents lead to fast depolarization of neurons, weak attractors, short dwell times and rapid transitions between attractors. Some basins of attractions are visited for such a short time that no action may follow, no chance for other neuronal groups to synchronize, it is hard to identify such states.

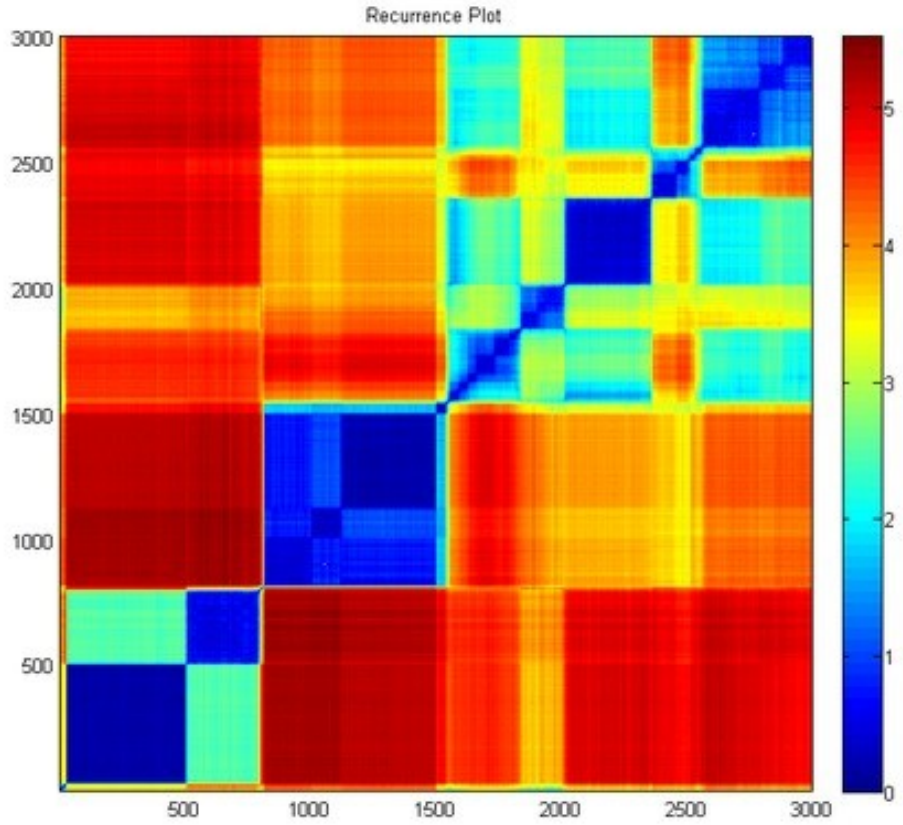
Psychological perspective: short attention span, hard to concentrate, learn new associations, feeling of confusion and fleeting thoughts.

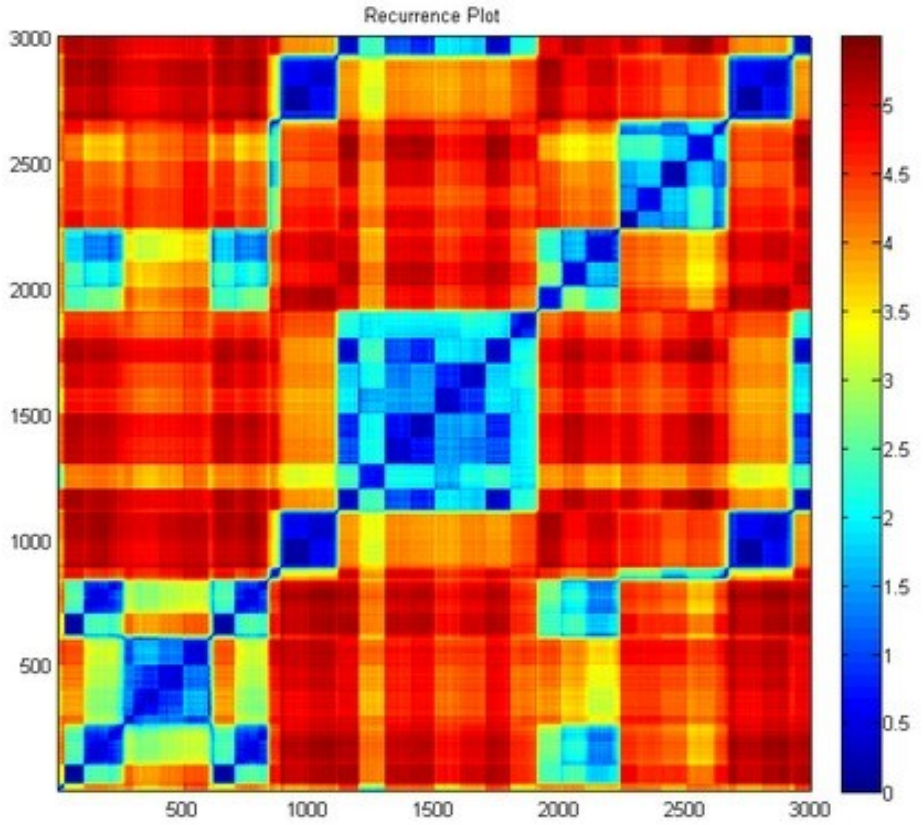
ASD-Normal-ADHD spectrum

b_inc_dt = 0.005 *b_inc_dt* = 0.01 *b_inc_dt* = 0.02



Simulations of rapid stimulation in autism





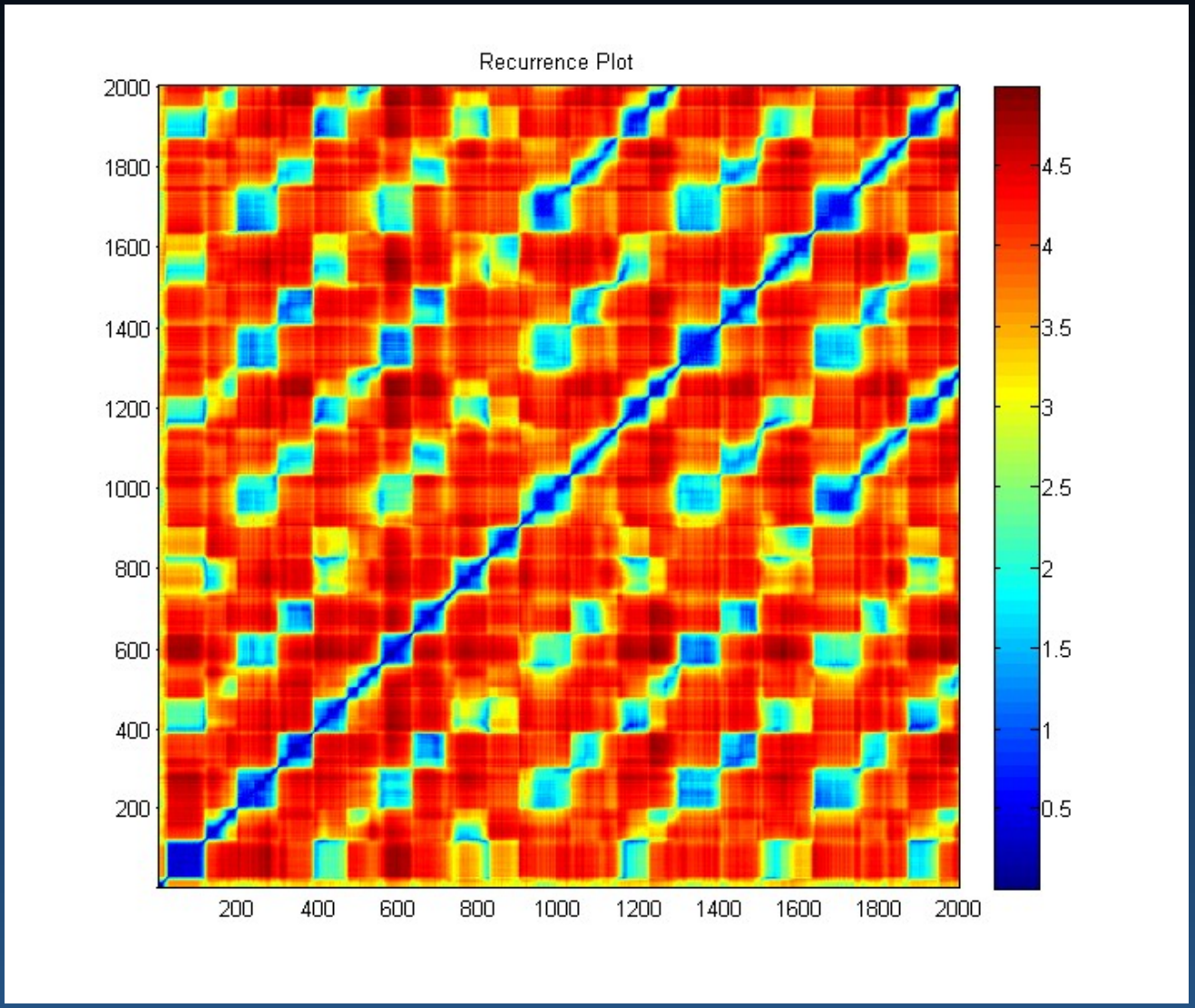
Normal presentation speed:
skipping some words,

no associations

Fast presentation speed:
more internal states

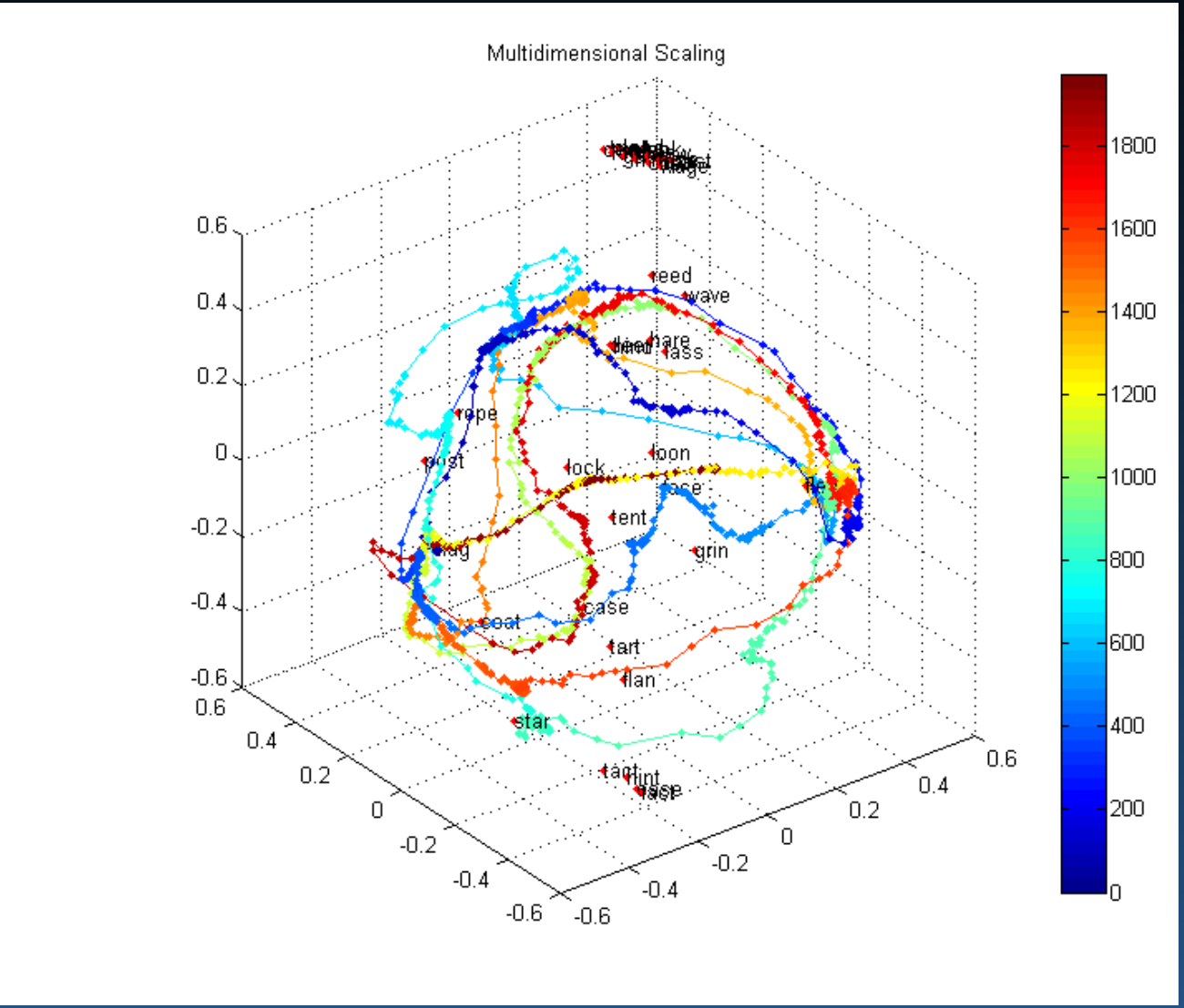
some associations arise

Reccurence plot: long trajectories

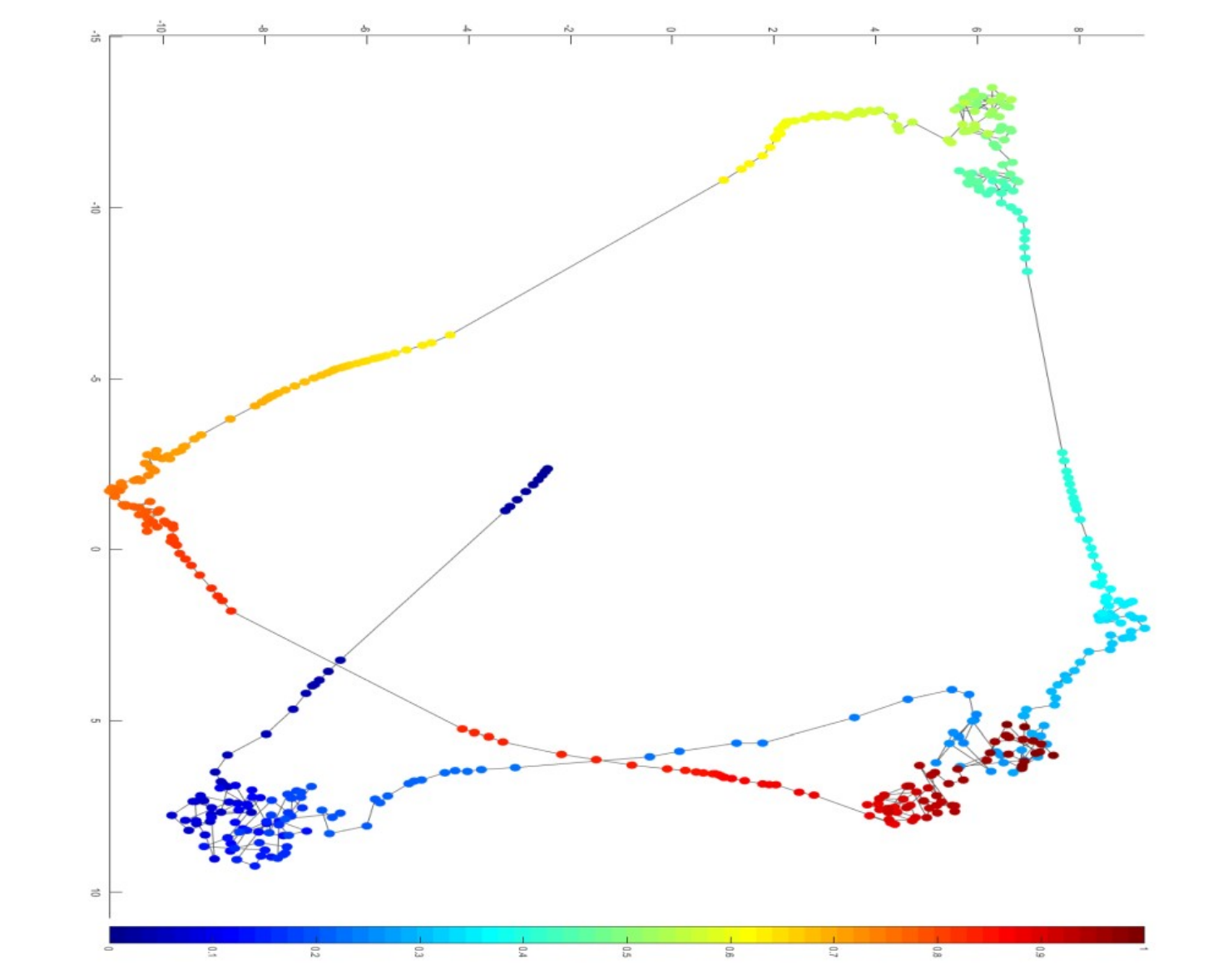


Recurrence plot in 40-words microdomain, starting with the word “flag”. The system tends to come back to similar pattern.

MDS: long trajectories



MDS visualization in 40-words microdomain, starting with the word “flag”. Several patterns close to the initial pattern, showing strong priming effects.



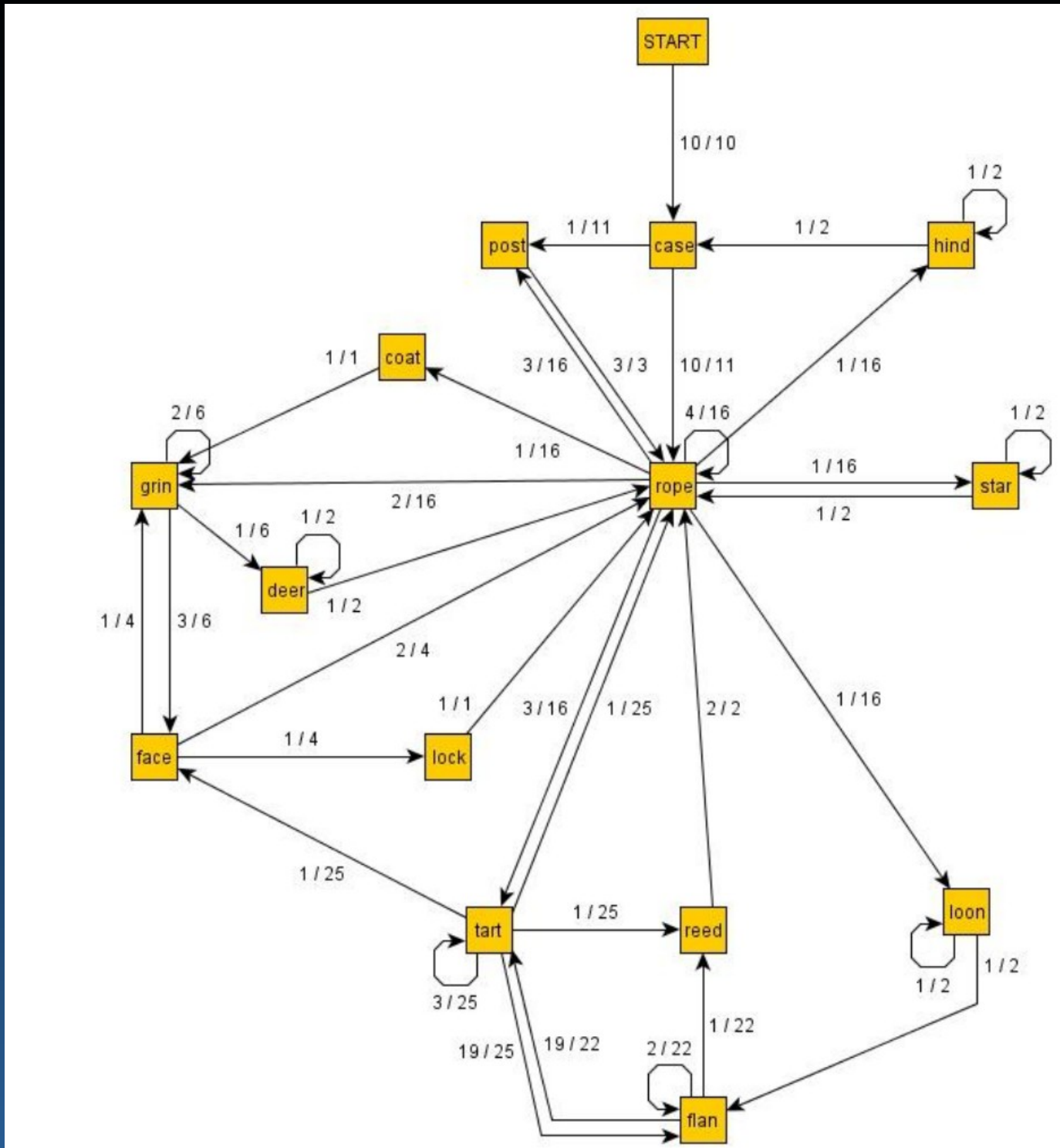
Trajectory visualization using Stochastic Neighbor Embedding (tSNE) technique shows basins of attractors that trap neurodynamics for a short time.

Graph of transitions between attractors, average of the 10 runs.

Why these particular transitions?

Escape from attractors makes some micro-features (neural units) stronger and some weaker, but visualization does not show details.

The whole landscape of available attractors is very dynamic! Transition probabilities change, dimensions (features) are re-scaled, and basins of attractors vanish.



Conspiracy will get you.
Ride the Hale-Bopp comet!

Conspiracies in our brains

How people start to believe in conspiracies?

The soul selects its own society, then shuts the door (Emily Dickinson).

Slow and rapid scenarios are possible, here only rapid presented:

- Emotional situations => neurotransmitters => neuroplasticity => fast learning, must be important.
- Fast learning => high probability of wrong interpretation.
- Traumatic experiences, hopelessness, decrease brain plasticity and leave only strongest association – strongly connected pathways.
- Conspiracy theories form around such associations, "frozen" pathways lead to brain activations forming strong attractors, distorting rational thinking.

- Such strong associations save brain energy and cannot be changed by rational arguments, that influence weaker associations only.
- This explanation becomes so obviously obvious …

Model: concept vectors derived from a corpus + MDS or Growing Neural Gas visualization (Martinetz & Schulten, 1991).

Self-organized networks

Rapid freezing of high neuroplasticity (RFHN) can be implemented using various neural network simulators. DemoGNG used here implements competitive learning models.

Dot = brain activation pattern. Links = associations between patterns.

Network nodes should calculate similarity of input data to their parameters.
Input vector **X** is compared to node parameters **W.** Similar = minimal distance.
Competition: find node *j=c* with **W** most similar to **X**.
Cooperation: nodes on a grid close to the winner *c* should behave similarly. Define the "neighborhood function" *O*(*c*):

$$h(r, r_c, t) = h_0(t) \exp\left(-\|r - r_c\|^2 / {\sigma_c}^2(t)\right)$$

$$\text{For } \forall i \in O(c)$$

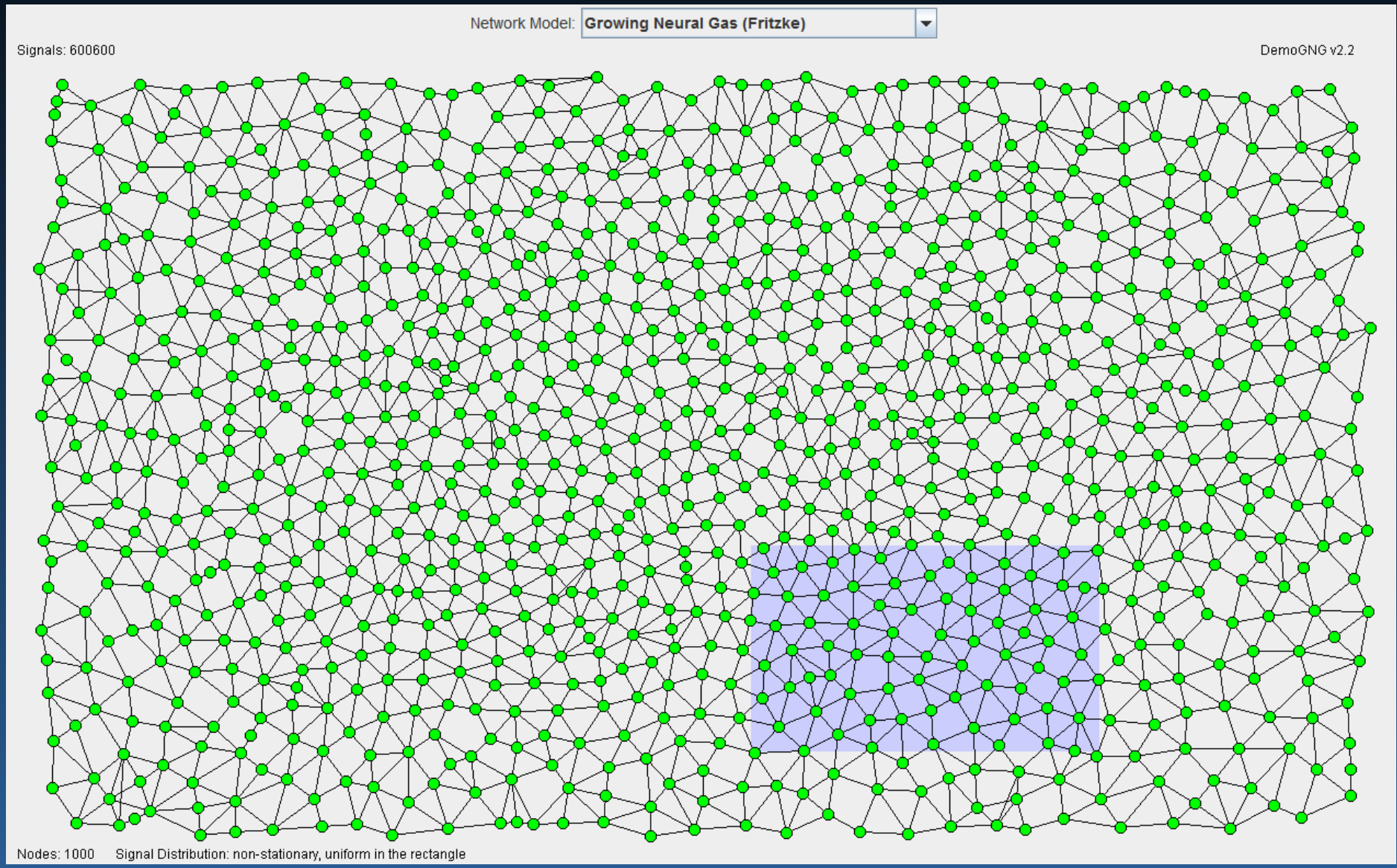
$$\mathbf{W}^{(i)}(t+1) = \mathbf{W}^{(i)}(t) + h(r_i, r_c, t)\left[\mathbf{X}(t) - \mathbf{W}^{(i)}(t)\right]$$

Node number **c** is most similar to the input vector **X**
It is a winner, and it will learn to be more similar to **X**, hence this is a "competitive learning" procedure.

Brain: those neurons that react to some signals pick it up and learn.

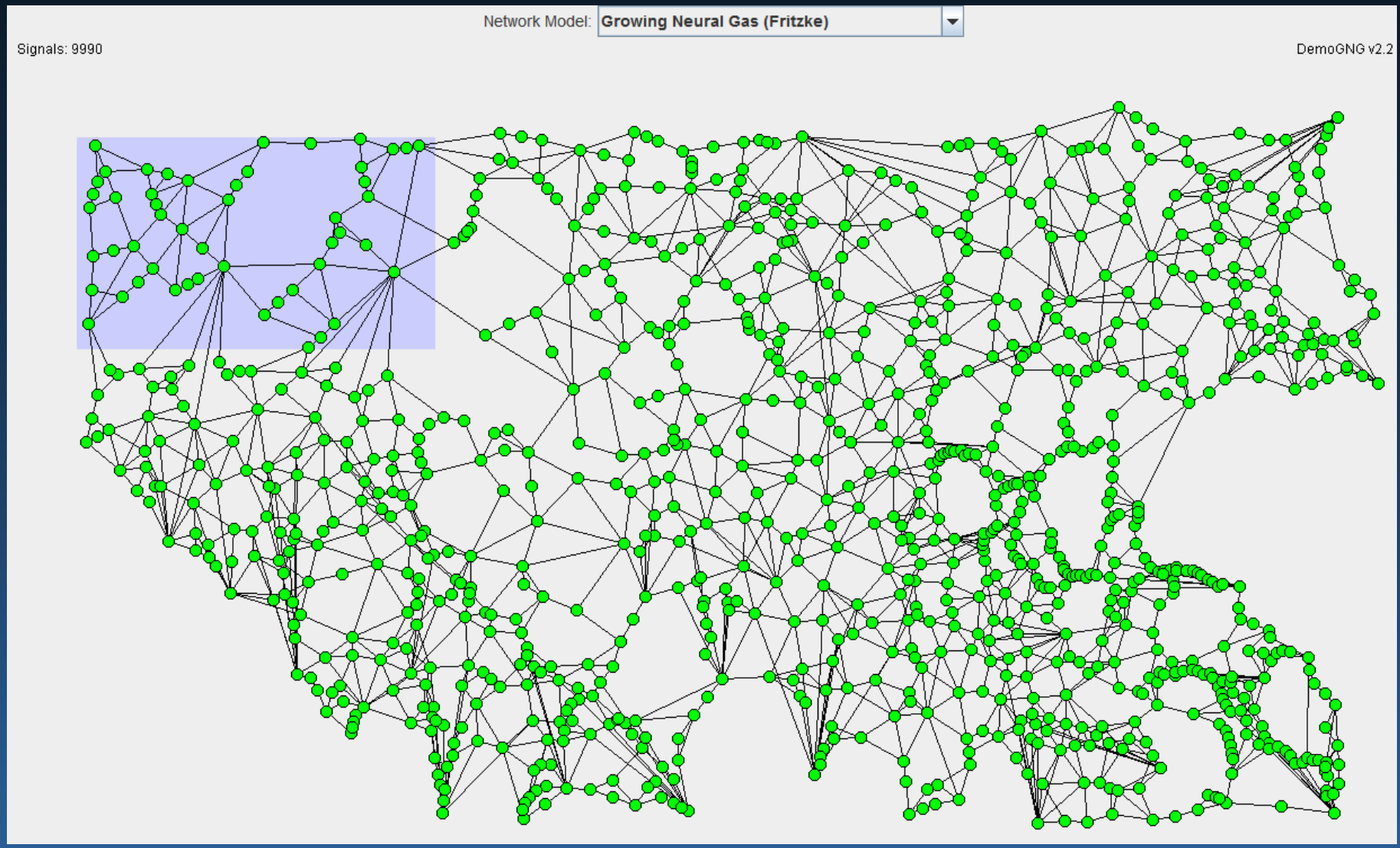
Internalization of environment

Episodes are remembered and serve as reference points, if observations are unbiased they reflect reality, creating correct associations and realistic evaluation.



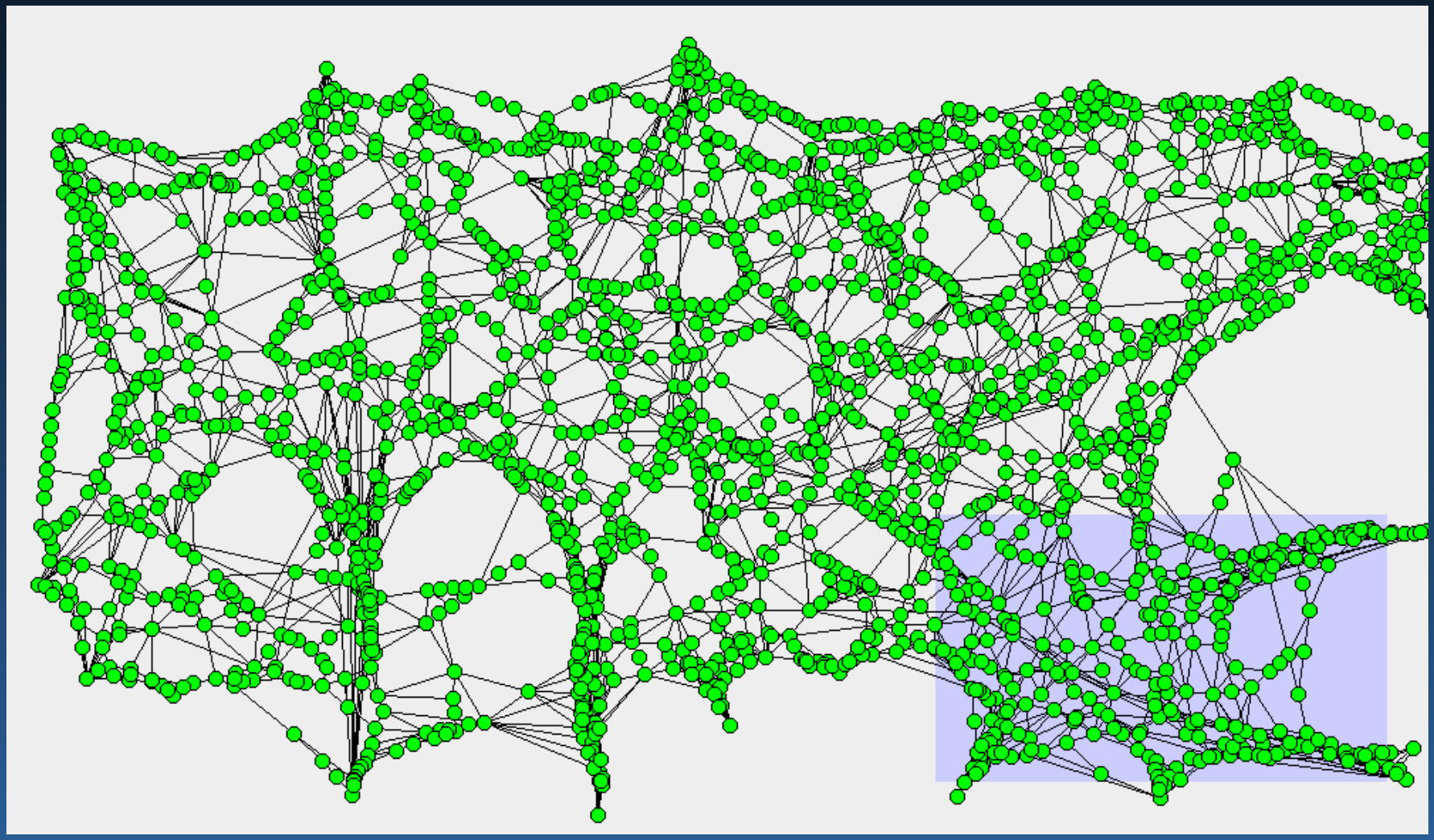
Extreme plasticity

Brain plasticity (learning) is increased if strong emotions are involved.
Rapid learning is not accurate, and if it is followed by depressive mood it leads to severe distortions, false associations, simplistic understanding.



Conspiracy views

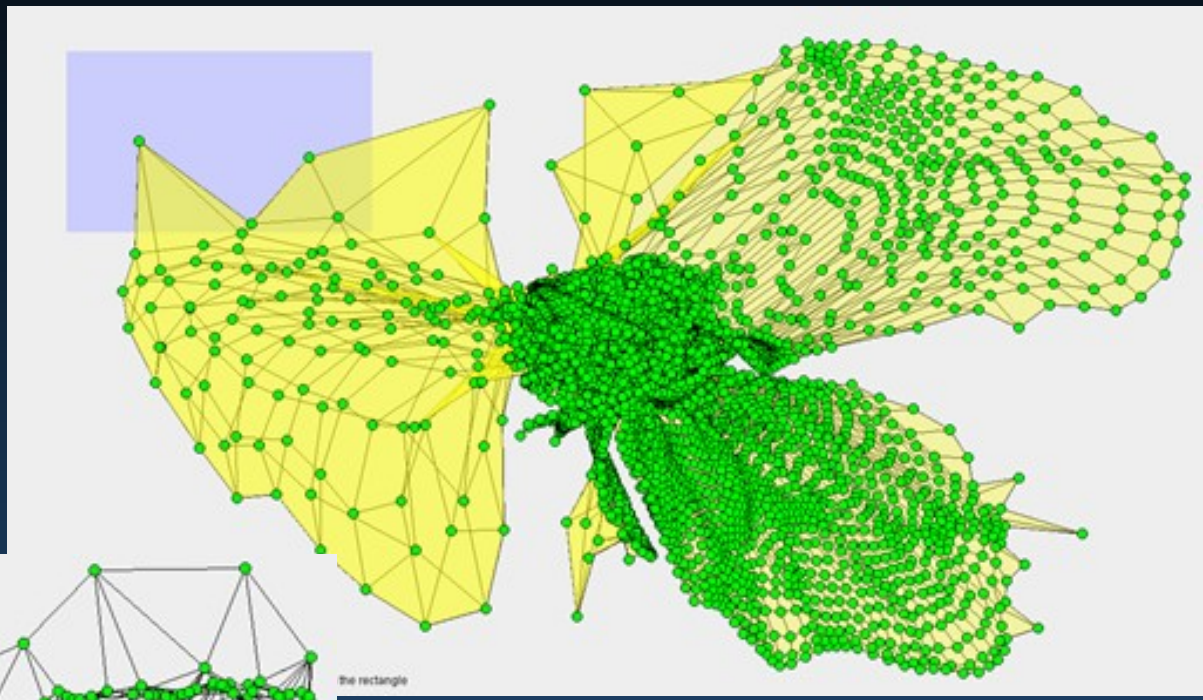
Illuminati, masons, Jews, UFOs, or twisted view of the world leaves big holes and admits simple explanations that save mental energy, creating „sinks” that attract many unrelated episodes.

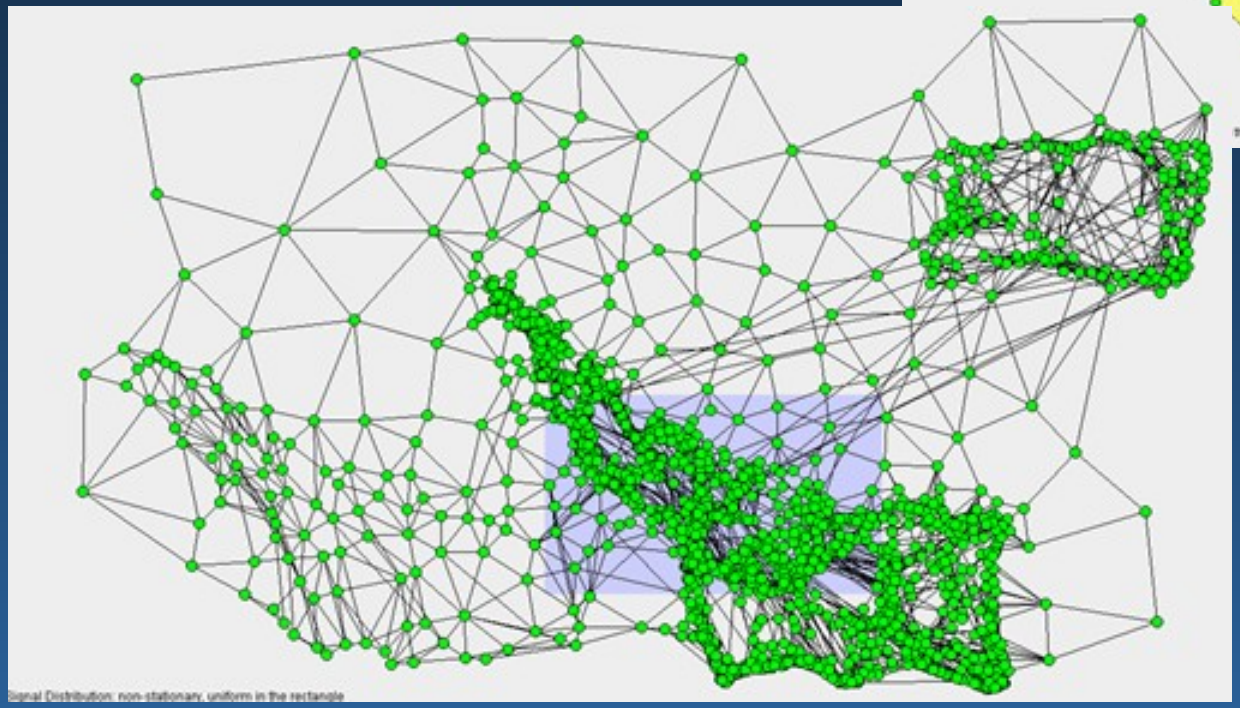


Memoids …

Totally distorted world view, mental processes are reduced to a memplex …

Ready to sacrifice oneself for a great idea.

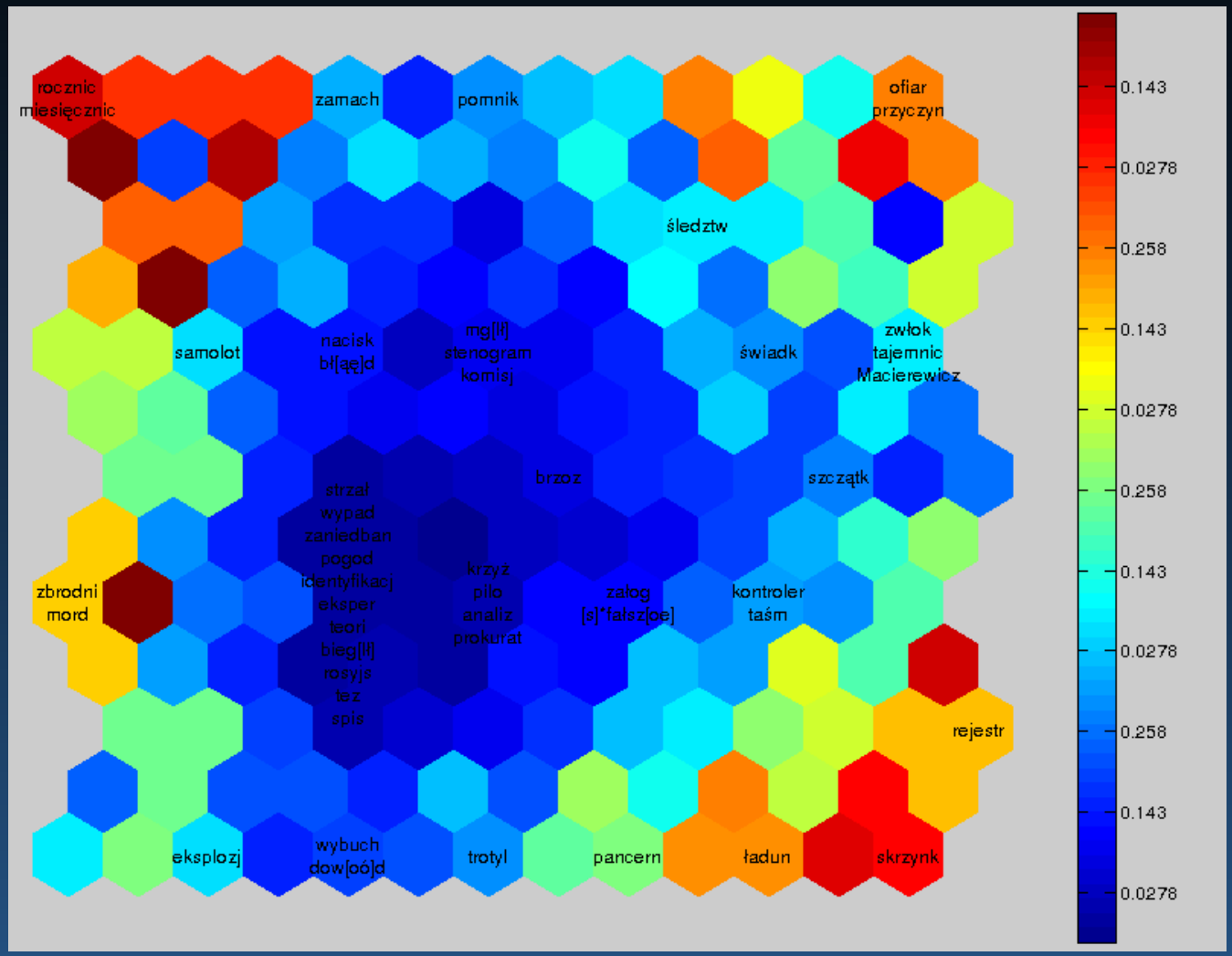




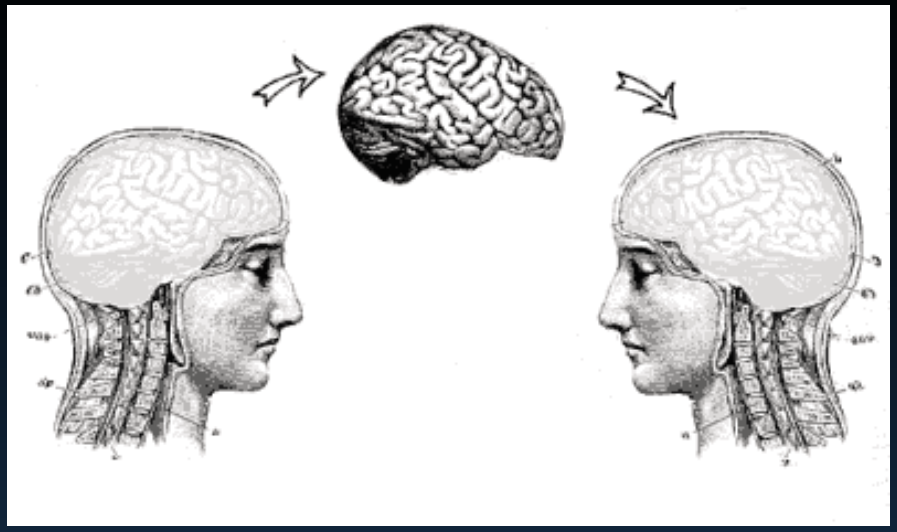
SOM on real newspaper data

Different groups of people read different newspapers, are exposed to different media and social networks, resulting in different network of concepts and sharp polarization of opinions.

Big sinks attract neurodynamics manifesting in strong automatic associations with core concepts.

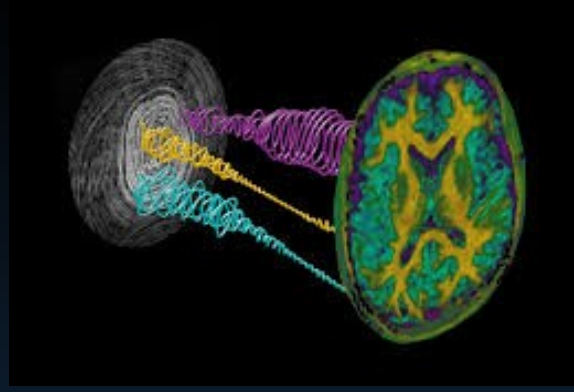


Different associative networks make communication almost impossible.
Work in progress (with J. Szymanski et al.)



Direct brain-brain transfers.

Sharing concepts



Communication requires shared space of concepts.
Human brains have roughly similar structure, but diverse experiences lead to idiosyncratic encoding of information in fine-scale functional topographies (J.V. Haxby papers).

Since 1986 in the USA and UK Core Knowledge Foundation tries to define common cultural codes, from preschool to the end of primary school, in precise sequence. Their motto: knowledge builds on knowledge.

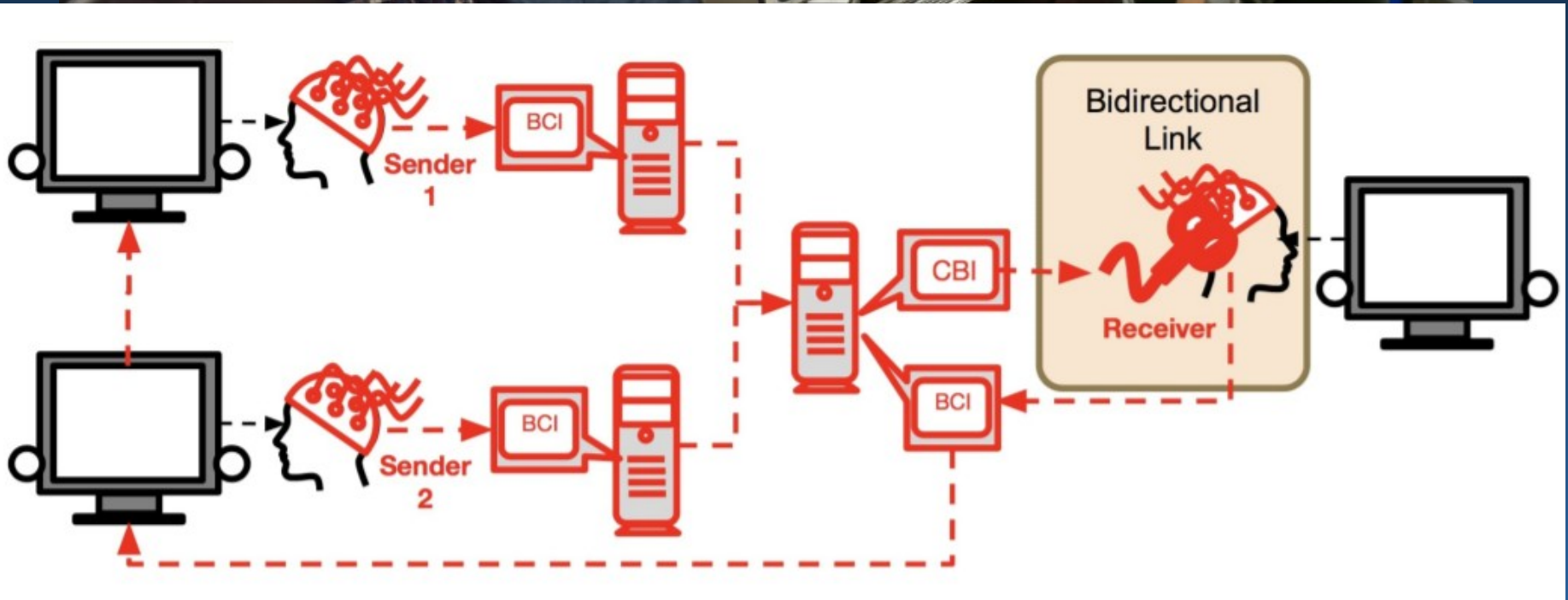
Communication space: how do we understand concepts, associate them, how my conceptual network differs from yours, and how should I effectively communicate my views and ideas to you?

"Causal structure of some systems cannot be fully captured by even the most detailed microscale description". **Causal emergence** is a new concept in analysis of communication processes. At the macroscale equivalent information in different brains may be activated at the symbolic level, but at the microscale it is not possible.

Varley, T., & Hoel, E. (2021). Emergence as the conversion of information: A unifying theory. ArXiv:2104.13368
Guntupalli et al., A Model of Representational Spaces in Human Cortex. Cerebral Cortex.
Base on hyperalignment to define common model space for the whole cortex.

Thought transfer?

Reading brain states => transforming => recreating in another brain.

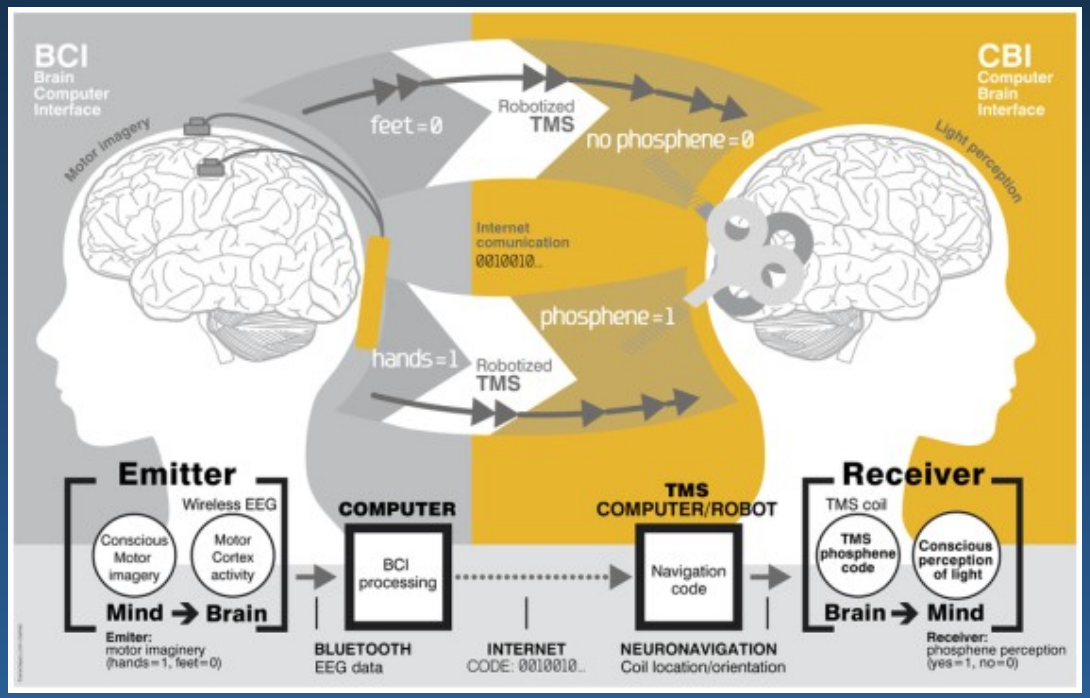


Direct brain activation

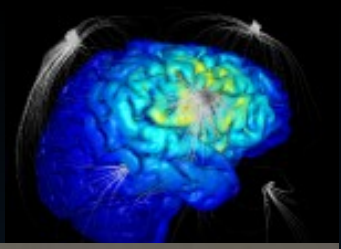
Sony has patent for direct streaming of multimedia to the brain.

Method and system for generating sensory data onto the human neural cortex. US Patent 6536440 B1

It should enable “sensory experiences” by firing “pulses of ultrasound at the head to modify firing patterns in targeted parts of the brain.”
This would allow the device to trigger various senses, including taste and sound, and even allow the deaf to hear again.
Will it facilitate Brain-to-Brain Communication?

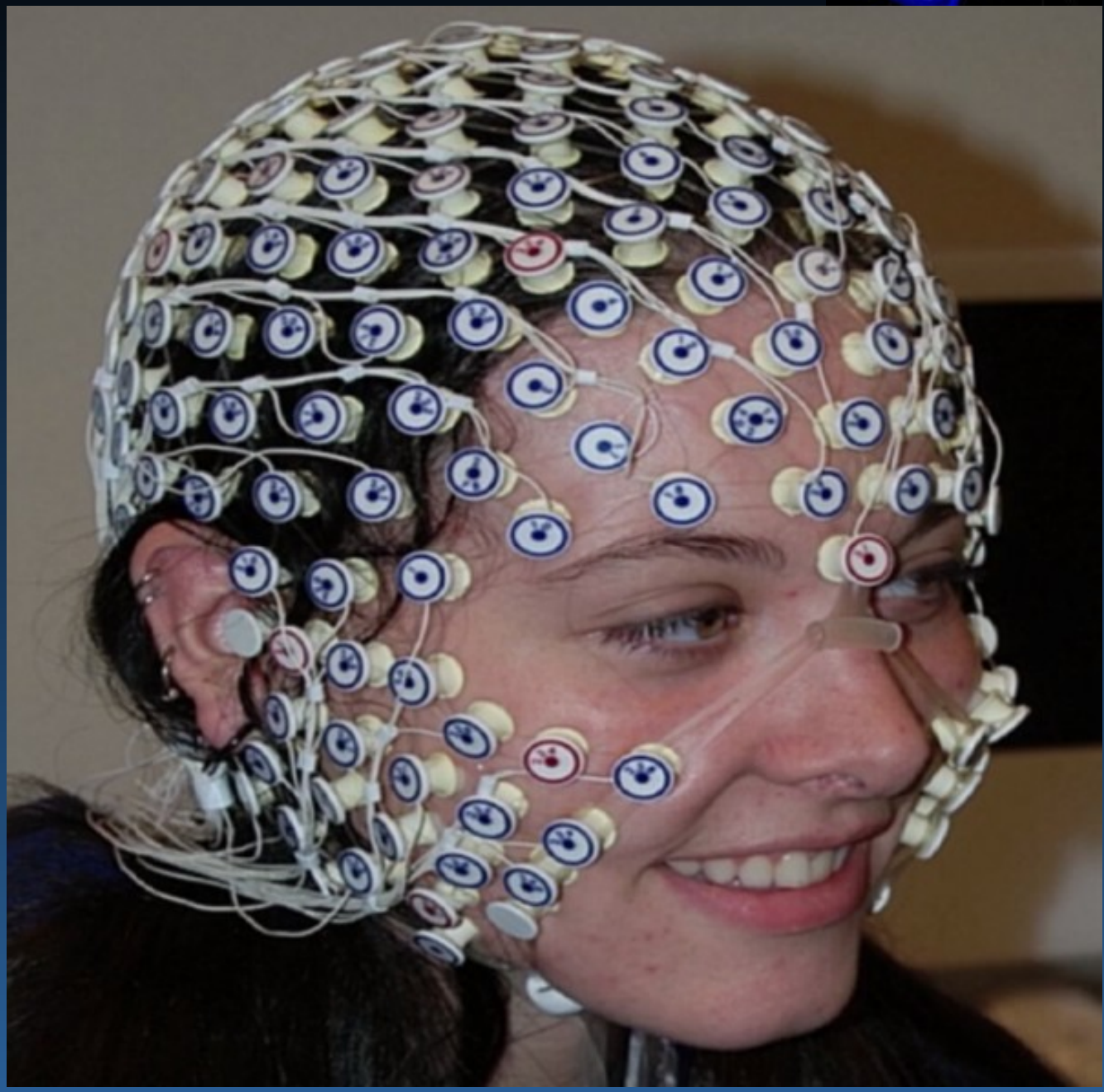


Resonance through HD DCS?



Reading brain states => transforming to common space => duplicating in other brains … Depression, neuro-plasticity, teaching!

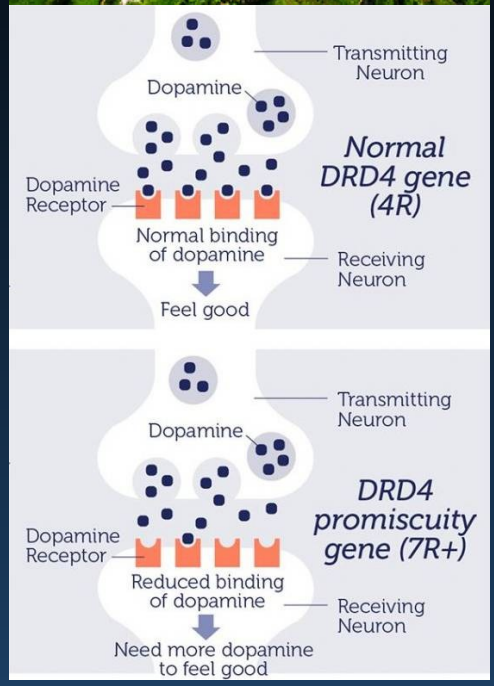
Multielectrode DCS stimulation with 256 electrodes induces changes in the brain increasing neuroplasticity.



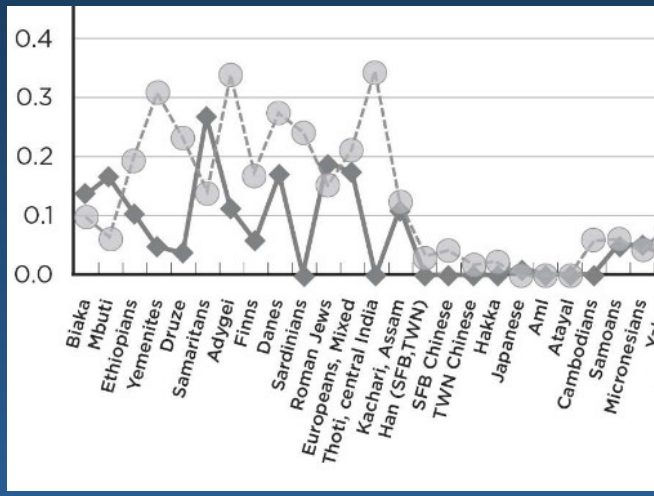
Genes and behavior

Study of links between ecology, food production, and culture show collectivist/individualist divisions in China, Bali, Turkey.

- Ecology ↔ living style ↔ food production ↔ culture ↔ genes.
- Rice was domesticated ≈ 10 000 years ago, requires massive communal work: transform the ecosystem: terraces, irrigation, harvesting, dividing up water fairly. In Bali it is regulated by priests in water temples. Irrigation system in Dujiuangyan (Western China) build 2000 years ago has > 5 000 km^2.
- In some parts of China farming is individual, people grow wheat. In standard tests those people are like Westerners. They also show high inventiveness—patent filings and higher rates of divorce.
- Dopamine D4 receptor is coded by extremely variable DRD4 gene. 25 human variants, controlling brain's reward system.



Most common: 4R variant, ½ of East Asians and Europeans.
7R variant, producing a receptor less responsive to dopamine in the cortex, associated with novelty seeking, extroversion, and impulsivity. It became much more common 10-20 000 years ago.
7R variant occurs in 20-30% of Europeans, and European Americans, but only in 1% of East Asians.



Some conspiracies are true ...

History's Greatest Lies (France. CBLBrights)

The bigger the lie, the more it will be believed!

- R. Nixon: Watergate 1972, illegal large scale invigilation, constant lies, cheating on taxes
- G.W. Bush and his administration: illegal war in Iraq, fabrication of evidence of WMD.
- D. Trump: lied about 30 000 times, defending his own self-image.
- B.L. Madoff: 2008 largest financial fraud of $65 billion piramid.
- F. Mitterand: sinking Greenpiece ship, Elysee Wire Tapping Scandal, long list of lies.
  J. Chirac lied to realize officers, who organized this terrorist attack.
- V. Putin: lies about nazi regime in Ukraine, lies about Chechen wars.

- Enron energy 100 B$, Ig Nobel Prize for "Most Creative Use of Imaginary Numbers".
- E. Holmes, Theranos 10 B$ company lied about diagnosis using small blood samples.

How can new believe anyone?

In search of the sources of brain's cognitive activity

Project „Symfonia 4”, 2016-22

FACULTY OF PHYSICS, ASTRONOMY AND INFORMATICS

CENTRE FOR MODERN INTERDISCIPLINARY TECHNOLOGIES

INSTITUTE OF PHYSIOLOGY AND PATHOLOGY OF HEARING

nencki institute of experimental biology

VIRTUAL BR41N.IO HACKATHON

April 17-18, 2021

during the

Spring School 2021*

IEEE brain g.tec

*BR41N.IO and Spring School 2021 are part of g.tec's Teaching Plan 2021 with more than 140 hours of online courses and lectures.

1. PLACE WINNER

"NeuroBeat"

BCI application

Team members: Alicja Wicher, Joanna Maria Zalewska, Weronika Sójka, Ivo John Krystian Derezinski, Krzystof Tołpa, Lukasz Furman, Slawomir Duda

IMPROVING HUMAN DAILY LIFE FUNCTIONING

NEUROHACKATOR 2021

21. - 23. MAY 2021 // ONLINE

FRIDAY
Organisers presentation

workshops with Judges

SATURDAY
Project development in groups

STARTS 10 a.m.

SUNDAY
Evaluation

ENDS 10 a.m.

working 24h

REQUIREMENTS:

1. Create a team consisting of **3-5 people**.
2. Fill in the Registration Form (available on Facebook event).

DO YOU HAVE ANY QUESTIONS?

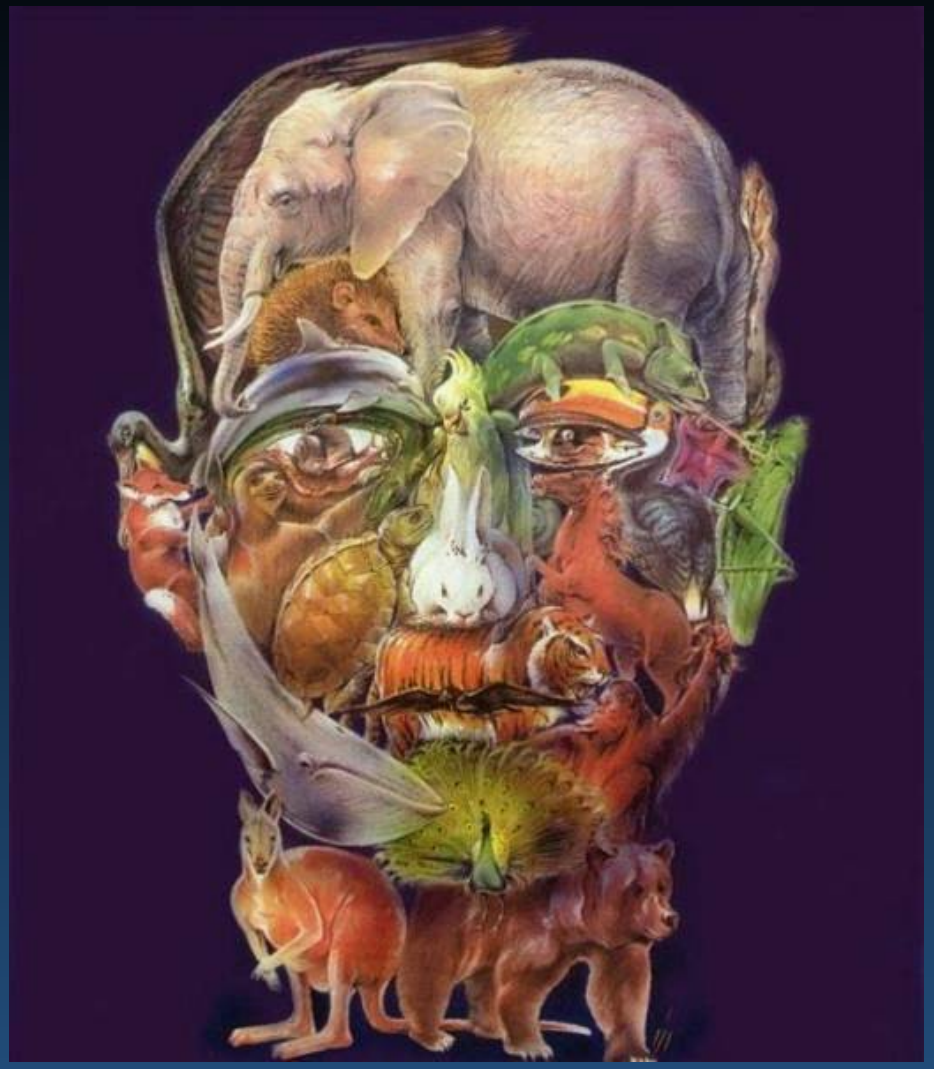
Write an e-mail:
NEUROTECHTOR@GMAIL.COM

Neurotechnology Scientific Club
Center for Modern Interdisciplinary Technologies
at Nicolaus Copernicus University in Toruń
Wileńska 4 Street

Thank you for synchronizing your neurons!

Our Neuroinformatics and Artificial Intelligence group, a part of the Dynamics, Mathematical Analysis and Artificial Intelligence Center, has announced new positions for PhD students, postdocs and visiting profs from abroad!
Please join our efforts!

Info is on our DAMSI webpage

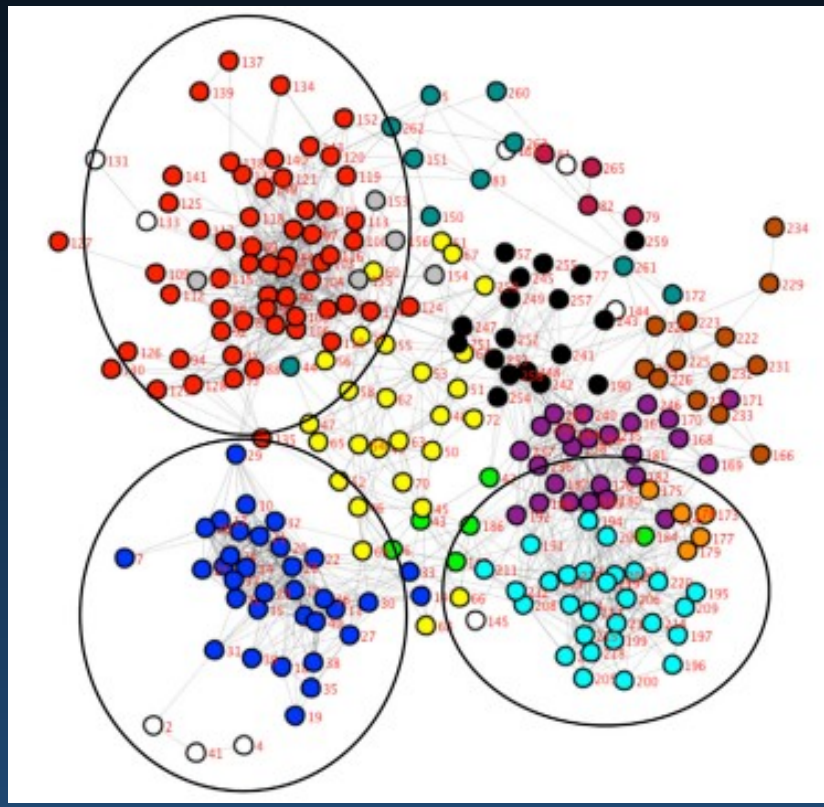


Google: Wlodek Duch
=> talks, papers, lectures, Flipboard …

Common model space

How to compare activations of different brains to bring them into resonance?

Guntupalli et al, A Model of Representational Spaces in Human Cortex. Cerebral Cortex (2016).

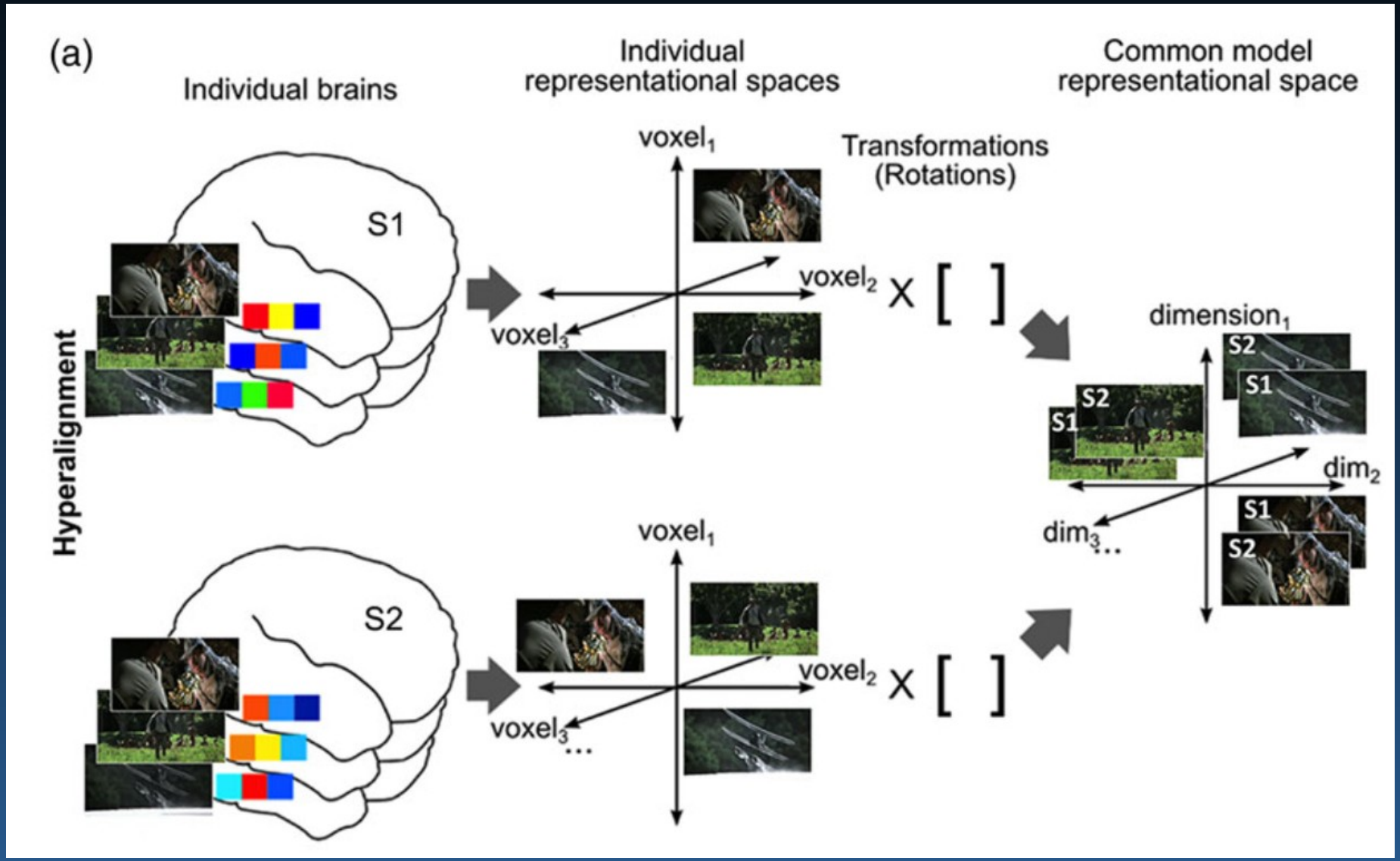


"Searchlight hyperalignment" algorithm to define common model space for the whole cortex.
The model aligns representations of complex, dynamic stimuli across brains in occipital, temporal, parietal, and prefrontal cortices,
as shown by between-subject multivariate pattern classification and intersubject correlation of representational geometry, indicating that structural principles for shared neural representations apply across widely divergent domains of information.

The model provides a rigorous account for individual variability of well-known coarse-scale topographies, such as retinotopy and category selectivity, and goes further to account for fine-scale patterns that are multiplexed with coarse-scale topographies and carry finer distinctions.

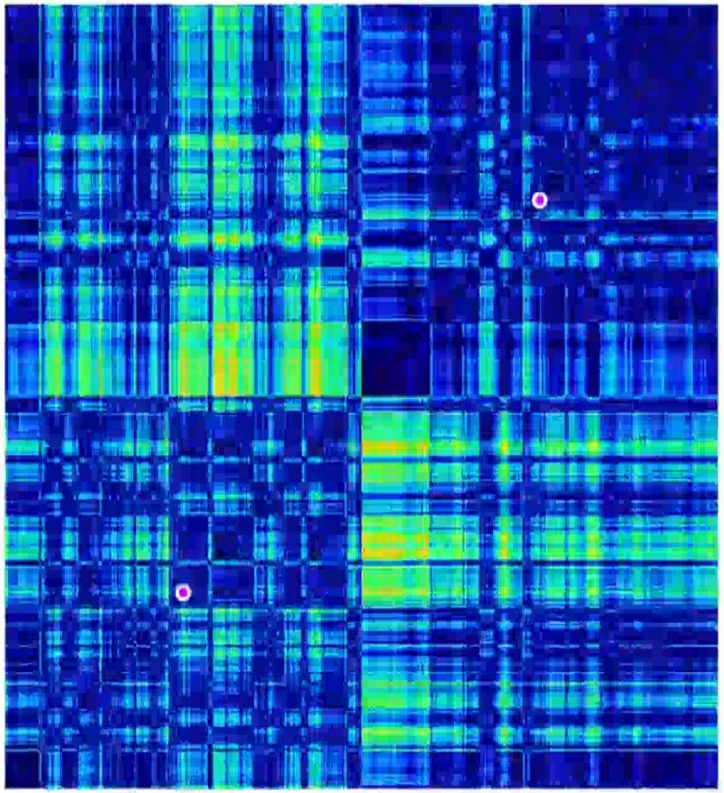
Hyperalignment

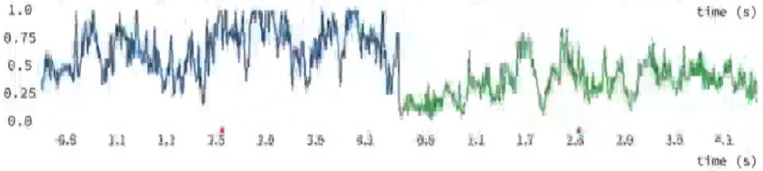
How to compare activations of different brains to bring them into resonance?

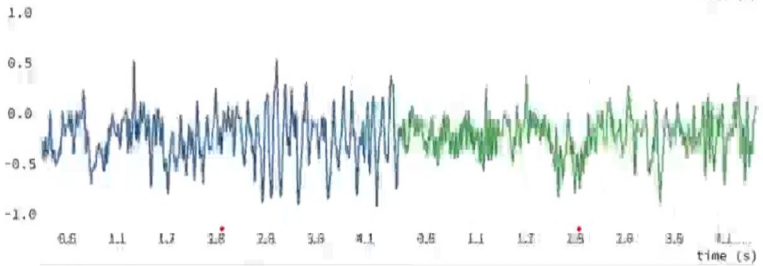


/perform*

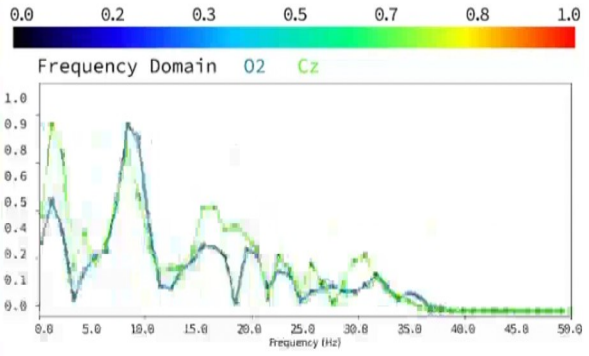
wavelets | E1 Param | E2 Param | set STFT | set Dist | colormap | set RP | set SSIM | vis SSIM

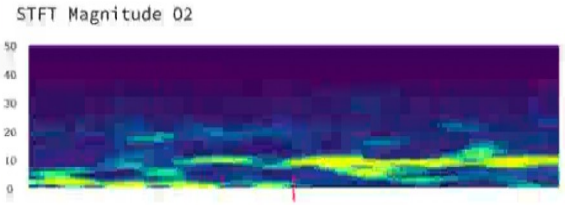


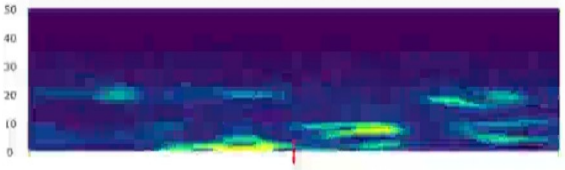


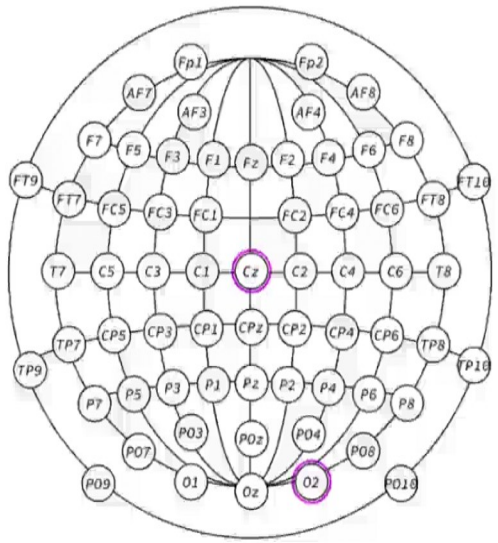


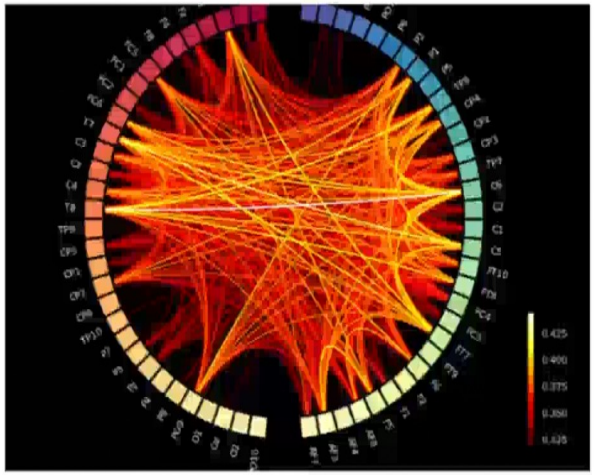
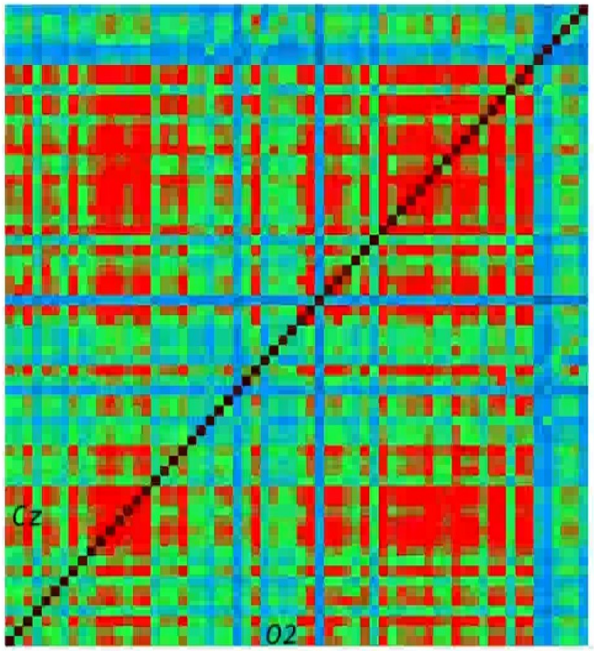
Init | Play









########## PYRQA COMPUTATION FINISHED ##########

Structural similarity