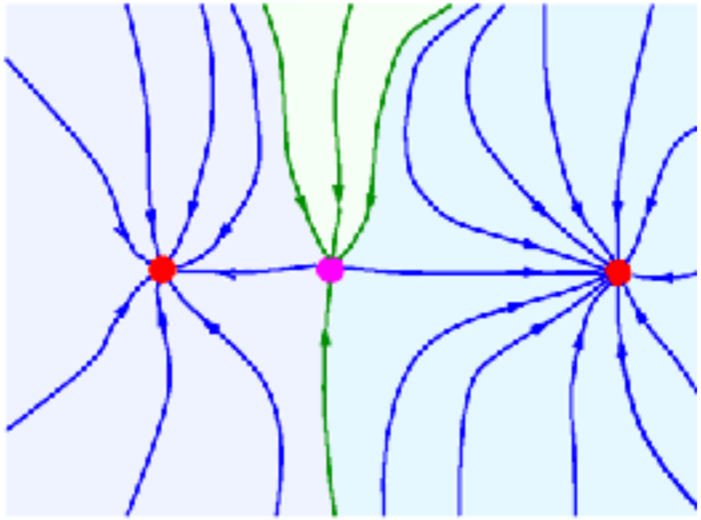


Schizophrenia, Attractors and Working Memory

Peggy Seriès, IANC
Informatics, University of Edinburgh, UK

pseries@inf.ed.ac.uk

CCN Lecture 5

This week's lectures

• Working memory models as example of dynamical models and attractor dynamics in biological neural networks;

• Application to Computational Psychiatry - Schizophrenia.

• Reading: Chapter 3 of Computational Psychiatry Primer.

1) What is schizophrenia?
2) What is working memory?
3) Attractor Paradigm for working memory/ delayed activity
- Hopfield Net
- Back to the Ring Model
4) Application to schizophrenia

1) What is Schizophrenia ?

❖ Major neuropsychiatric disorder
❖ 1% lifetime risk
❖ can start as sudden psychotic episode usually in early adulthood or gradual withdrawal and decline.

1. Positive symptoms: (appearance of abnormal experiences - psychosis)

- Threatening voices
- Thought Echoing
- Visual Hallucination
- Delusions

https://www.youtube.com/watch?v=KYHVbLLO2bU

1) What is Schizophrenia ?

2. Disorganisation:

• Pressure of thought

“ in my mind there ran like and endless clockwork a compulsive torturing uninterrupted chain of ideas”

• Incongruent behaviour

“ I really would like to slap your face, people like you are usually called SOBs”

• disturbed sense of self (“mirror sign”)

1) What is Schizophrenia ?

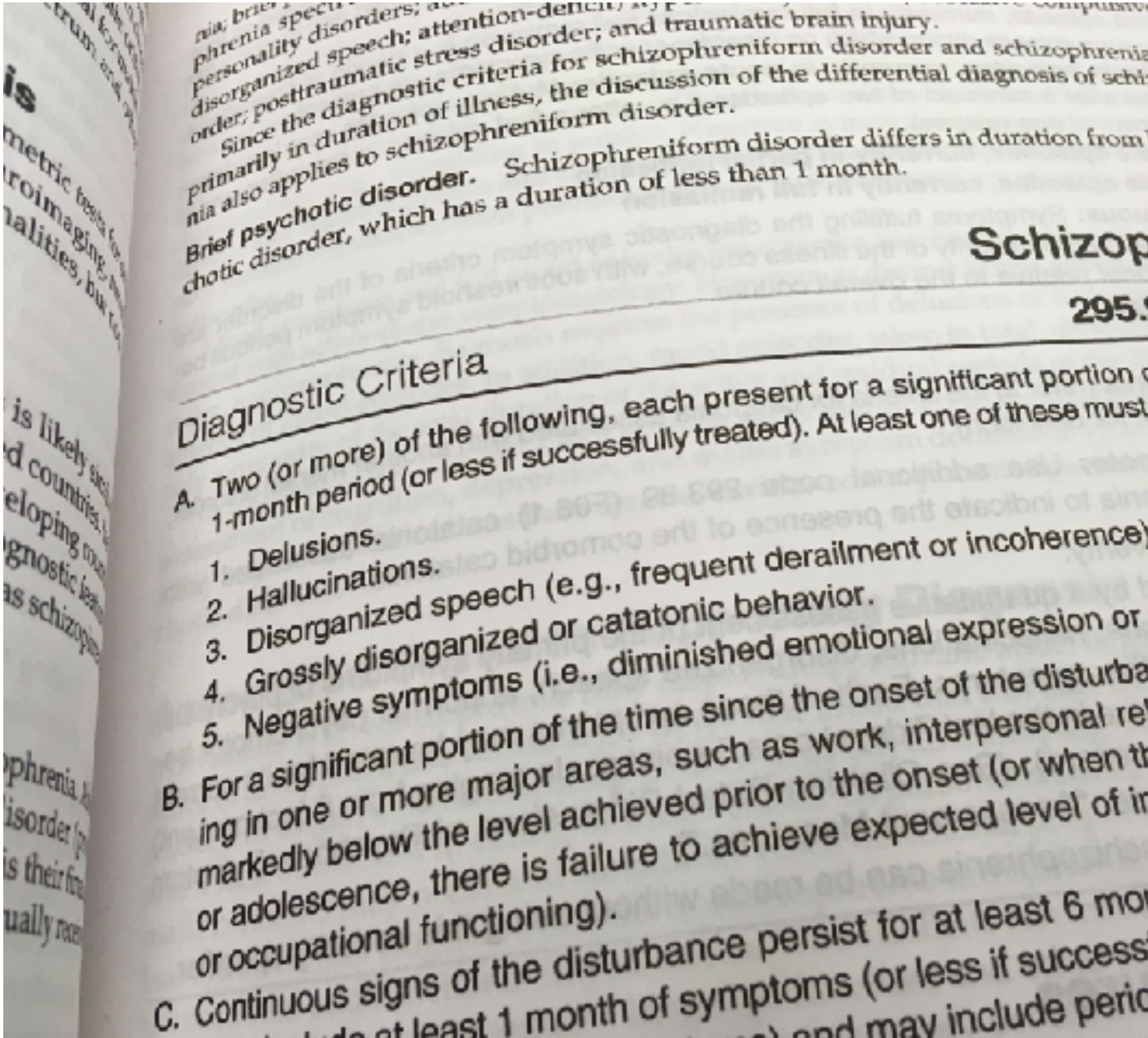
3. Negative Symptoms:
(deficits in normal experience)

- Poverty of thought / perseverance
- Loss of Initiative
- anhedonia: loss of pleasure
- social withdrawal

- **Cognitive deficits**

Difficulty in maintaining context, in learning associations.
working memory deficits

1) What is Schizophrenia ?



personality disorders; ... disorganized speech; attention-deficit/hyperactivity ... compulsive disorder; posttraumatic stress disorder; and traumatic brain injury.

Since the diagnostic criteria for schizophreniform disorder and schizophrenia differ primarily in duration of illness, the discussion of the differential diagnosis of schizophrenia also applies to schizophreniform disorder.

Brief psychotic disorder. Schizophreniform disorder differs in duration from brief psychotic disorder, which has a duration of less than 1 month.

Schizophrenia

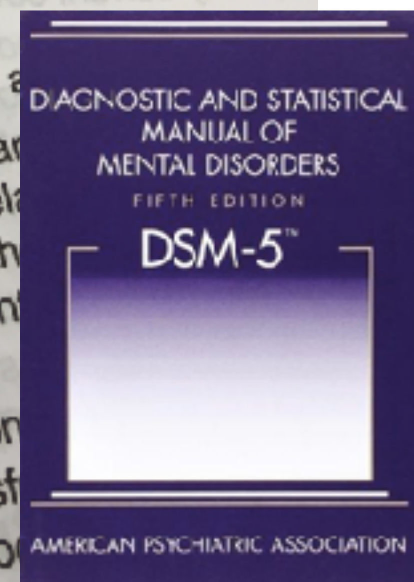
295.90 (F20.9)

Diagnostic Criteria

A. Two (or more) of the following, each present for a significant portion of time during a 1-month period (or less if successfully treated). At least one of these must be (1), (2), or (3):
   1. Delusions.
   2. Hallucinations.
   3. Disorganized speech (e.g., frequent derailment or incoherence).
   4. Grossly disorganized or catatonic behavior.
   5. Negative symptoms (i.e., diminished emotional expression or a

B. For a significant portion of the time since the onset of the disturba ing in one or more major areas, such as work, interpersonal rel markedly below the level achieved prior to the onset (or when th or adolescence, there is failure to achieve expected level of in or occupational functioning).

C. Continuous signs of the disturbance persist for at least 6 mon ... at least 1 month of symptoms (or less if successf ... and may include perio



1) What is Schizophrenia ?

- Cause is unknown : Combination of genetic and environmental factors
- No cure. Medication alleviate the (+) symptoms somewhat.

- **Disconnection hypothesis**
- **Dopamine hypothesis:** misfiring of dopaminergic neurons.

Antipsychotic drugs work on the dopamine system.

Amphetamines (DA agonist) can mimic some aspect of the disorder.

- **NMDA hypothesis:** Impairments in NMDA transmission.

Ketamine (NMDA receptor antagonist) mimic aspects of the disorder.

- Positive symptoms are most fascinating but negative/ cognitive symptoms create most suffering and disability.
- 10% suicide rate.

RDOC framework in Psychiatry

National Institute of Mental Health

- Building upon ongoing advances in the behavioural and neurobiological sciences
- "RDoC is an attempt to create a new kind of taxonomy for mental disorders by bringing the power of modern research approaches in genetics, neuroscience, and behavioral science to the problem of mental illness." [Elevag et al 2016]
- From a categorical to a dimensional approach.

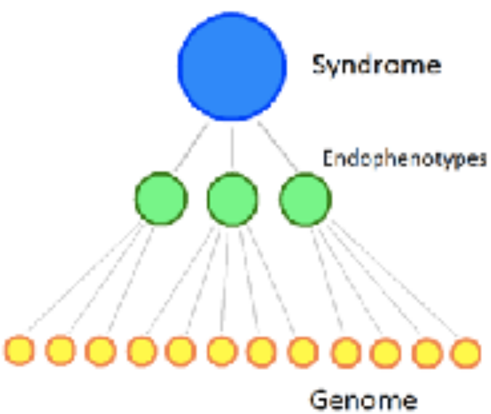
<table>
<tr><th colspan="10">NIMH Research Domain Criteria (RDoC)</th></tr>
<tr><th colspan="10">Functional Domains</th></tr>
<tr>
<td colspan="2">Negative Valence Systems
(e.g., fear, anxiety, loss)</td>
<td colspan="2">Positive Valence Systems
(e.g., reward, learning, habit)</td>
<td colspan="2">Cognitive Systems
(e.g., attention, perception, memory)</td>
<td colspan="2">Systems for Social Processes
(e.g., attachment, communication, perception of self & others)</td>
<td colspan="2">Arousal and Regulatory Systems
(e.g., arousal, circadian rhythms)</td>
</tr>
<tr><th colspan="10">Units of Analysis</th></tr>
<tr>
<td>Genes</td><td>Molecules</td><td>Cells</td><td>Circuits</td><td>Physiology</td><td>Behavior</td><td>Self-Reports</td><td>Paradigms</td><td>Genes</td><td>Molecules</td>
</tr>
</table>

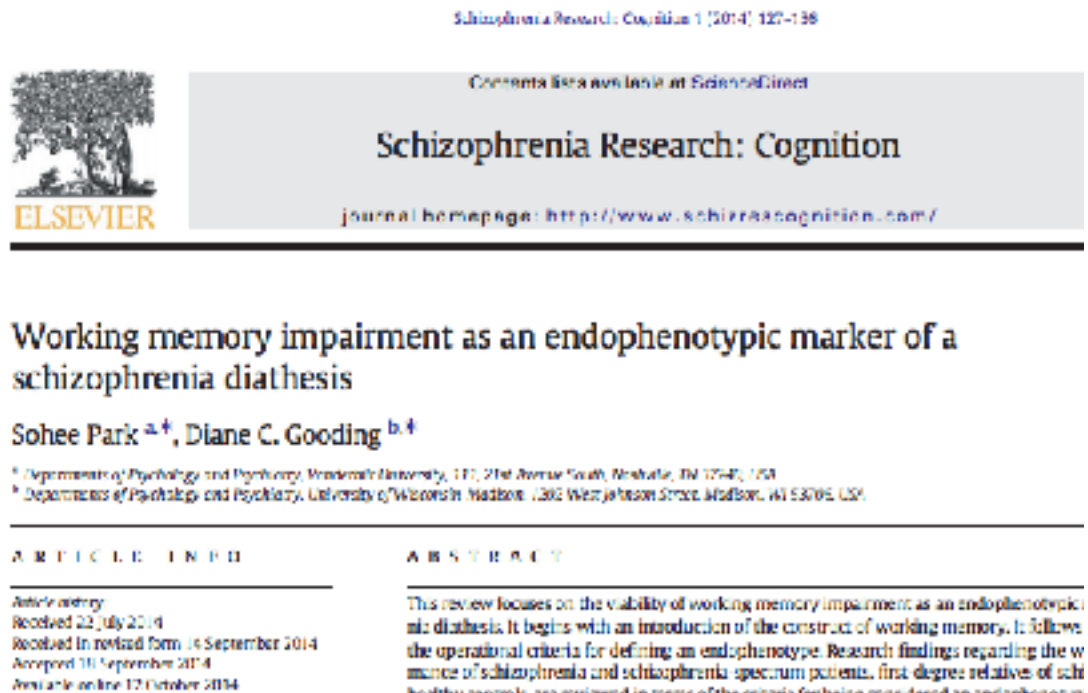
Linking Neurobiology and Psychiatry

- Working memory as a possible endophenotype?

endophenotype

n. a type of biological marker that is simpler to detect than genetic sequences and that may be useful in researching vulnerability to a wide range of psychological and neurological disorders. Endophenotypes may be a useful link between genetic sequences and their external emotional, cognitive, or behavioral manifestations. [coined in the 1970s by U.S. behavioral geneticist and clinical psychologist Irving I. Gottesman (1930–) and British psychiatrist James Shields (d. 1978)]




Schizophrenia Research: Cognition 1 (2014) 127–138
Contents lists available at ScienceDirect
Schizophrenia Research: Cognition
journal homepage: http://www.schizrescognition.com/

Working memory impairment as an endophenotypic marker of a schizophrenia diathesis

Sohee Park [a,*], Diane C. Gooding [b,*]

ARTICLE INFO

Article history:
Received 22 July 2014
Received in revised form 16 September 2014
Accepted 18 September 2014
Available online 12 October 2014

ABSTRACT

This review focuses on the viability of working memory impairment as an endophenotypic ... nic diathesis. It begins with an introduction of the construct of working memory. It follows ... the operational criteria for defining an endophenotype. Research findings regarding the w... mance of schizophrenia and schizophrenia spectrum patients, first degree relatives of schi...

2) What is working memory ? (a.k.a. short-term memory)

- The ability to hold information over a time scale of seconds to minutes
- a critical component of all cognitive functions (language, thoughts, planning etc..)

Delayed match-to sample task:
remember 'red'

- central to many learning disabilities and impaired in many mental disorders

http://news.bbc.co.uk/1/hi/health/4101431.stm

http://www.youtube.com/watch?v=M1Tui0Gbvq4&feature=related

2) What is working memory ?

Test yours!

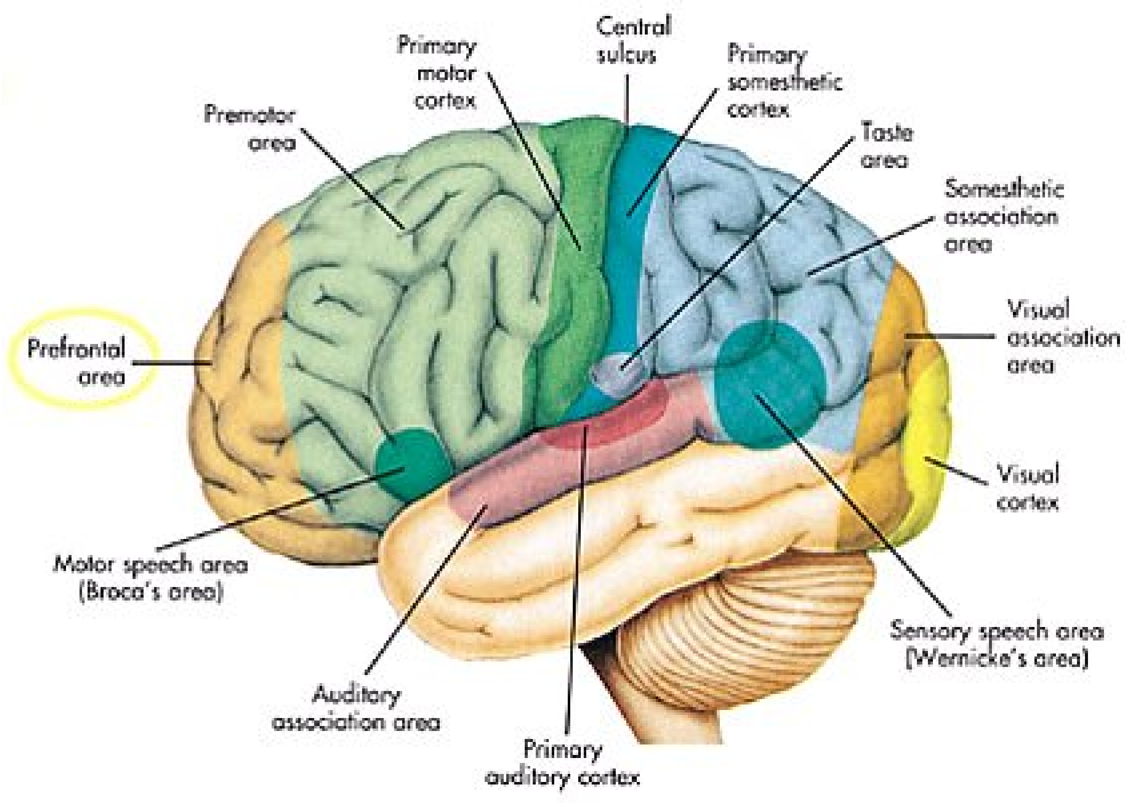
https://www.ted.com/talks/peter_doolittle_how_your_working_memory_makes_sense_of_the_world

2:21-3:49

TED Ideas worth spreading

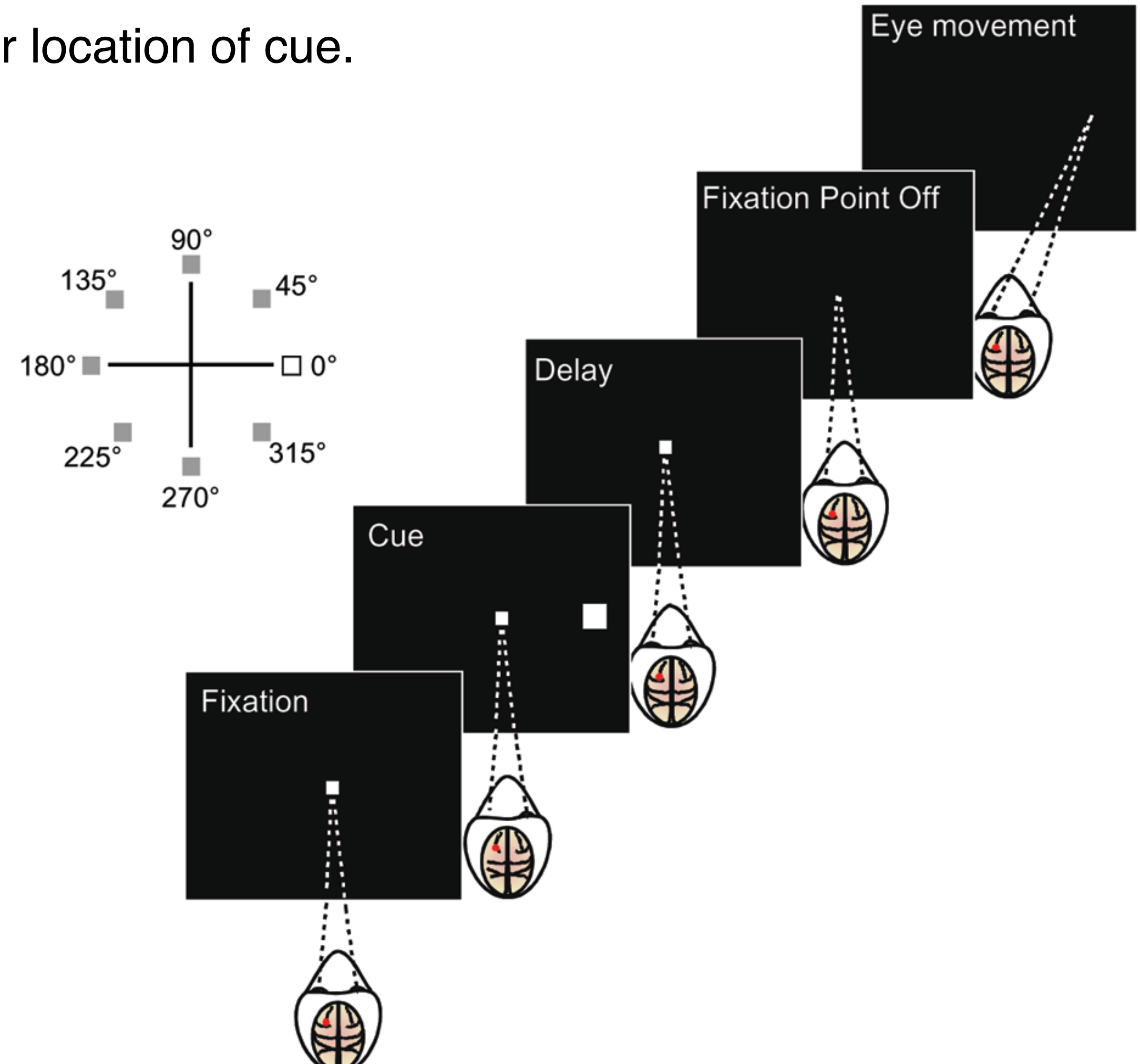
Sustained activity in PFC (1)

- Lesion and inactivation studies demonstrate crucial role of Prefrontal Cortex (PFC) in working memory, in particular dorsolateral PFC (PFdl).



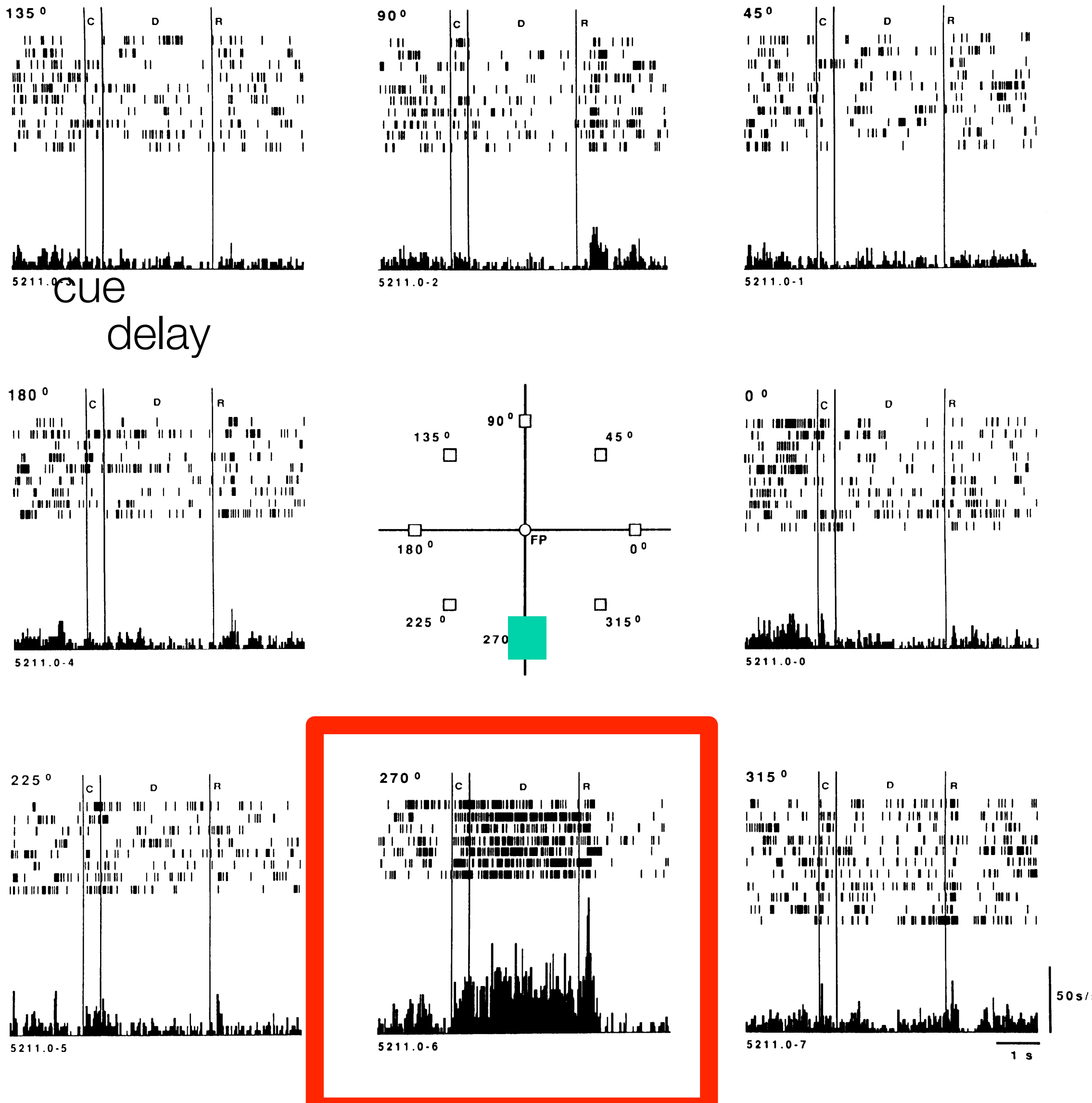
Sustained activity in PFC (2)

Oculo-motor delayed response task:
remember location of cue.



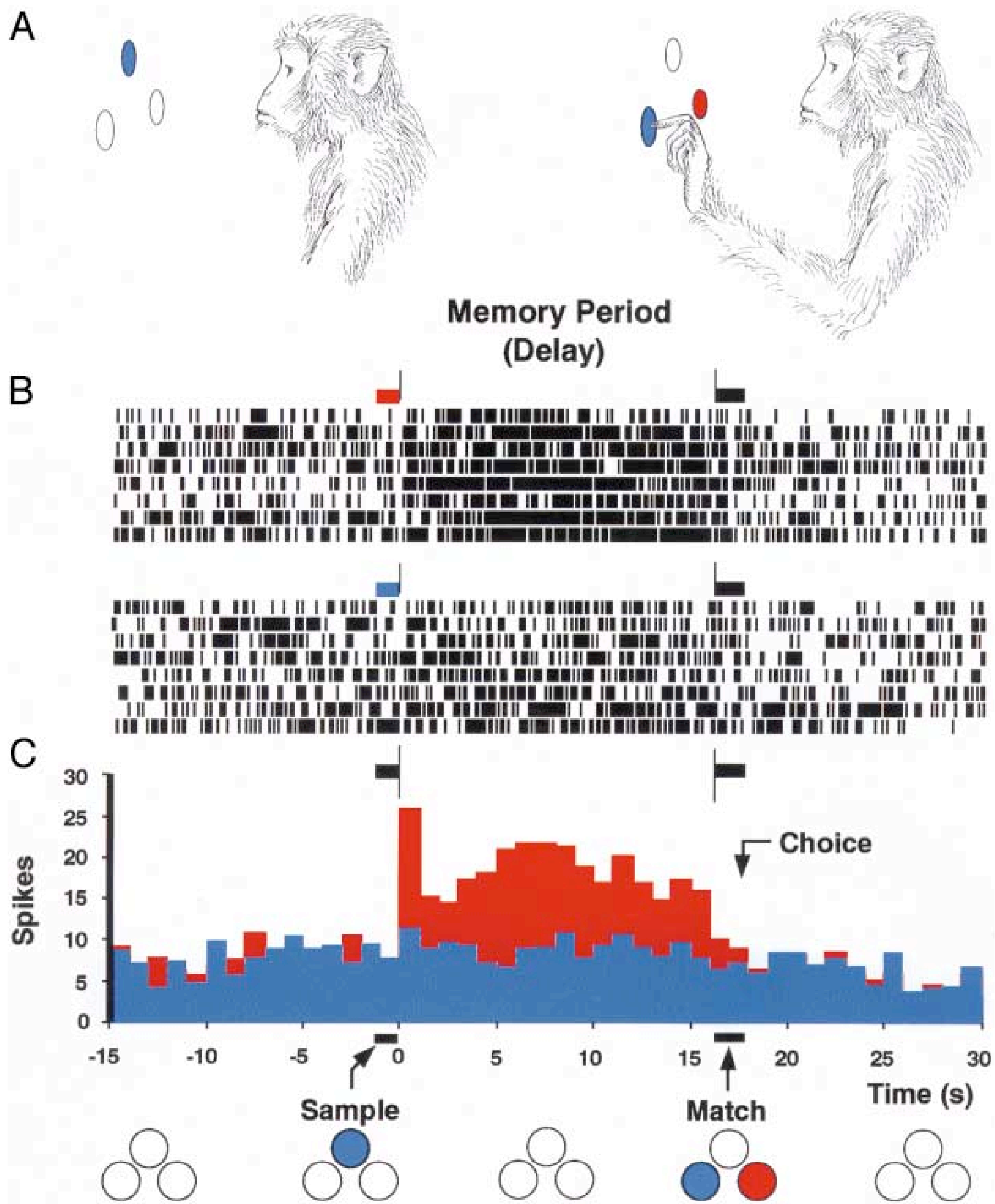
Sustained activity in PFC (2)

Patricia Goldman-Rakic
1937-2003



Funahashi et al, 1989

Sustained activity in inferotemporal Cx (IT)



Fuster and Jervey 1982

Working memory vs Long-term memory and Linking propositions

• Long-term memory : molecular or structural changes - connectivity between neurons.

• Short-term/ working memory: Dynamic process that has not yielded to molecular characterisation: "Sustained" a.k.a. "Delay" or "Persistent" Activity.

Towards a theory of Working Memory/ Sustained Activity

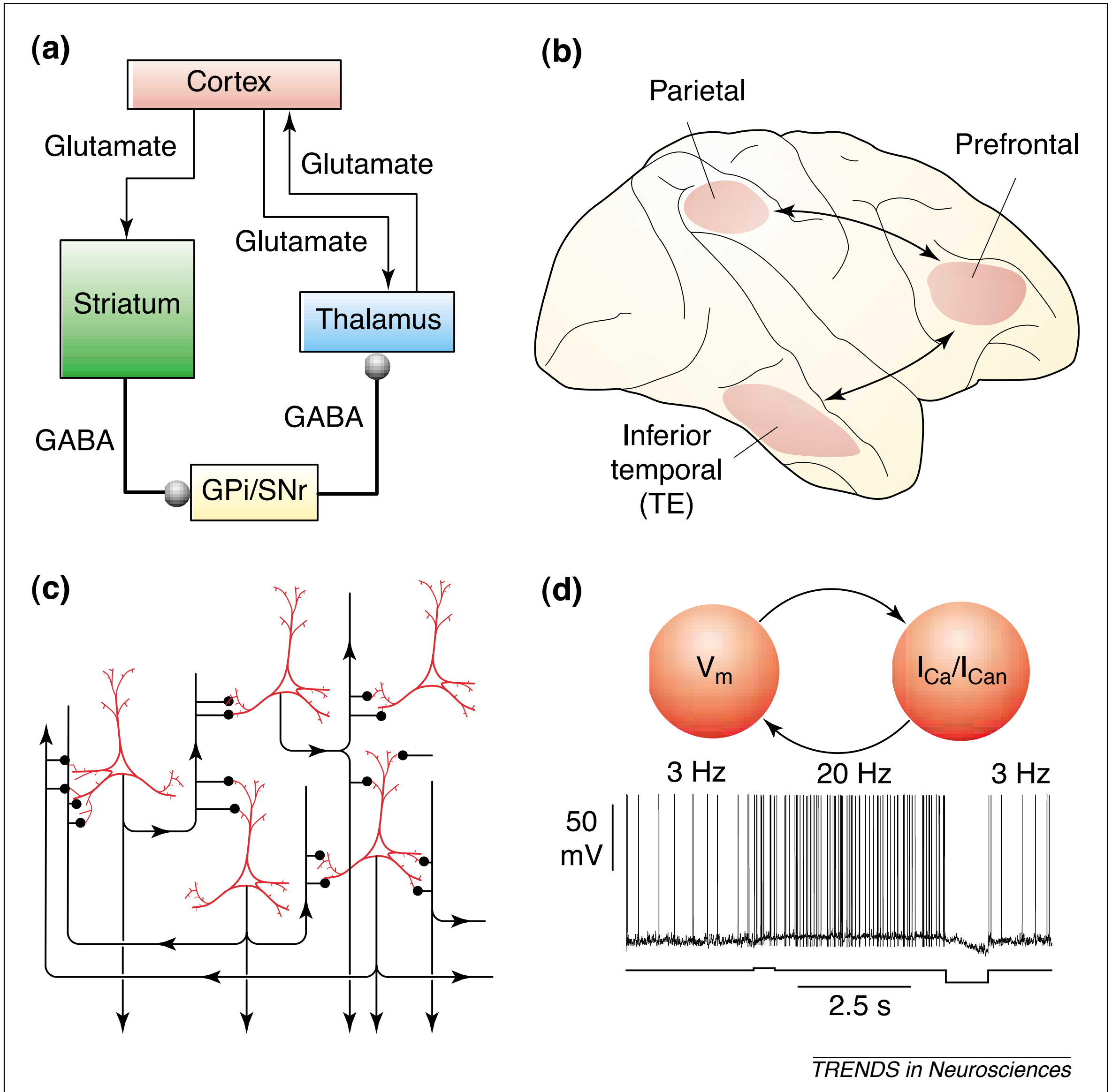
- A theory of working memory should answer:

- How it is initiated?
- Why does it persist ?
- What makes it specific?
- How does it end?

- Reason for capacity limit?
- Relationship with attention, long-term memory?

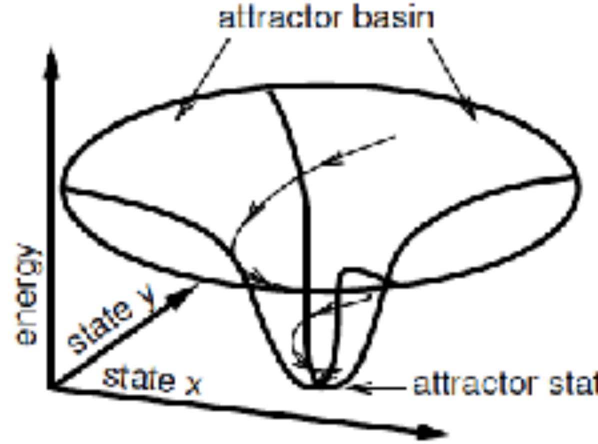
- Mechanism : reverberations through connections (which?), or cellular?
- Lots of experimental and theoretical work to answer these questions

Delay activity could be due to reverberations through…



3) Attractor Paradigm for Persistent Activity

- Since1970s it is proposed that delay activity patterns can be theoretically described by dynamical attractors' recurrent neural networks.

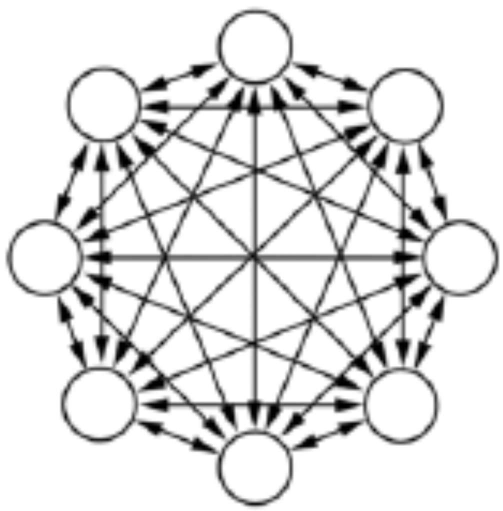


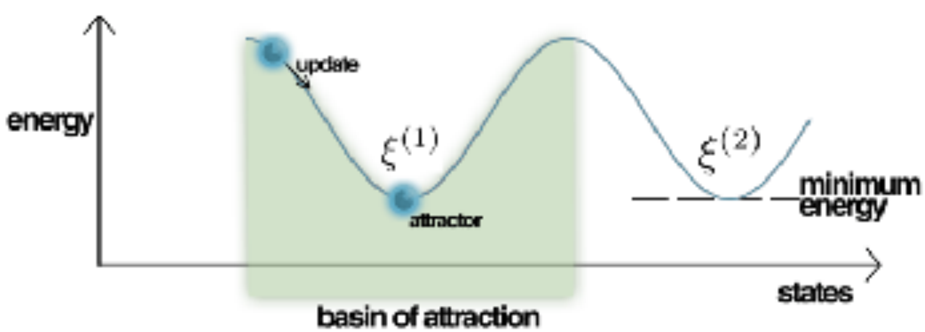
- Hopfield networks (John Hopfield (1982))

$$s_i = \begin{cases} 1 & \text{if } \sum_j w_{ij} s_j > \theta_i, \\ -1 & \text{otherwise.} \end{cases}$$

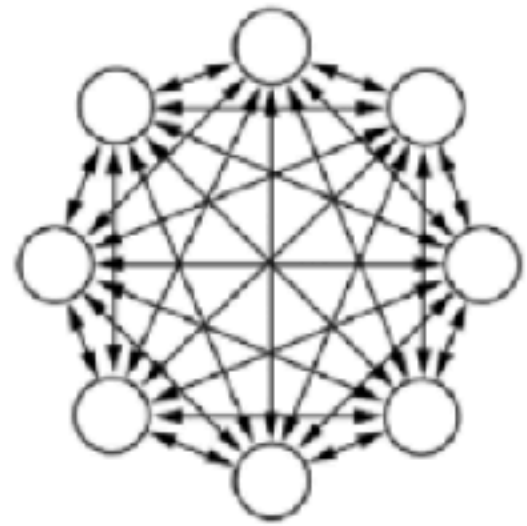
$$w_{ij} = \frac{1}{N} \sum_{k=1}^{k=N} \xi_i^k \xi_j^k$$

$$E = -\frac{1}{2} \sum_{i<j} w_{ij} s_i s_j + \sum_i \theta_i \ s_i$$





Hopfield Networks (1)



- A Hopfield net is a form of recurrent artificial neural network invented by John Hopfield (1982).
- Hopfield nets typically have binary (1/-1 or 1/0) threshold units:

$$s_i = \begin{cases} 1 & \text{if } \sum_j w_{ij} s_j > \theta_i, \\ -1 & \text{otherwise.} \end{cases}$$

where s_j state of unit j, and θ_i is the threshold

The weights have to follow: $w_{ii}=0$, $w_{ij}=w_{ji}$

- Hopfield nets have a scalar value associated with each state of the network referred to as the "energy", E, of the network, where:

$$E = -\frac{1}{2} \sum_{i<j} w_{ij} s_i s_j + \sum_i \theta_i \ s_i$$

Hopfield Networks (2)

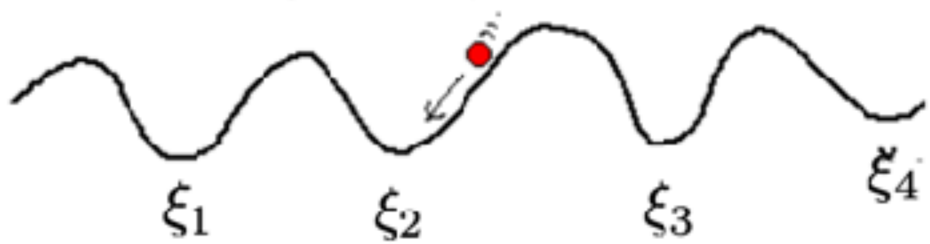
- Initialization is done by setting the values of the units to the start pattern (e.g. noisy or incomplete image).
- Running: at each step, pick a node at random and update (asynchronous update).

$$s_i = \begin{cases} 1 & \text{if } \sum_j w_{ij} s_j > \theta_i, \\ -1 & \text{otherwise.} \end{cases}$$

- Energy is guaranteed to go down & network to settle in local energy minima.
- Learning: the weights are set so as to 'shape' those local minima.

$$w_{ij} = \frac{1}{N} \sum_{k=1}^{k=N} \xi_i^k \xi_j^k$$

- Network will converge to stored state even if it is given only part of the state:



ξ_i =stored

Associative memories

• The Hopfield network is an associative/content addressable memory. It can be used to recover from a distorted input the trained state that is most similar to that input.
• e.g. if we train a Hopfield net with 5 units so that the state (1, 0, 1, 0, 1) is an energy minimum, and we give the network the state (1, 0, 0, 0, 1) it will converge to (1, 0, 1, 0, 1).

Reminiscent of human memory?

The Nobel Prize in Physics 2024

The Royal Swedish Academy of Sciences has decided to award the Nobel Prize in Physics 2024 to

John J. Hopfield
Princeton University, NJ, USA

Geoffrey Hinton
University of Toronto, Canada

"for foundational discoveries and inventions that enable machine learning with artificial neural networks"

They trained artificial neural networks using physics

This year's two Nobel Laureates in Physics have used tools from physics to develop methods that are the foundation of today's powerful machine learning. John Hopfield created an associative memory that can store and reconstruct images and other types of patterns in data. Geoffrey Hinton invented a method that can autonomously find properties in data, and so perform tasks such as identifying specific elements in pictures.

When we talk about artificial intelligence, we often mean machine learning using artificial neural networks. This technology was originally inspired by the structure of the brain. In an artificial neural network, the brain's neurons are represented by nodes that have different values. These nodes influence each other through connections that can be likened to synapses and which can be made stronger or weaker. The network is *trained*, for example by developing stronger connections between nodes with simultaneously high values. This year's laureates have conducted important work with artificial neural networks from the 1980s onward.

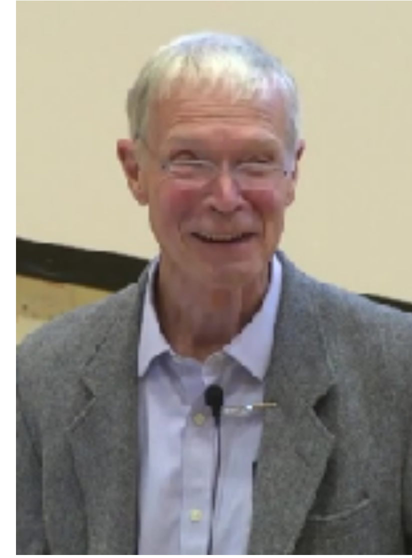
John Hopfield invented a network that uses a method for saving and recreating patterns. We can imagine the nodes as pixels. The *Hopfield network* utilises physics that describes a material's characteristics due to its atomic spin – a property that makes each atom a tiny magnet. The network as a whole is described in a manner equivalent to the energy in the spin system found in physics, and is trained by finding values for the connections between the nodes so that the saved images have low energy. When the Hopfield network is fed a distorted or incomplete image, it methodically works through the nodes and updates their values so the network's energy falls. The network thus works stepwise to find the saved image that is most like the imperfect one it was fed with.

Geoffrey Hinton used the Hopfield network as the foundation for a new network that uses a different method: the *Boltzmann machine*. This can learn to recognise characteristic elements in a given type of data. Hinton used tools from statistical physics, the science of systems built from many similar components. The machine is trained by feeding it examples that are very likely to arise when the machine is run. The Boltzmann machine can be used to classify images or create new examples of the type of pattern on which it was trained. Hinton has built upon this work, helping initiate the current explosive development of machine learning.

"The laureates' work has already been of the greatest benefit. In physics we use artificial neural networks in a vast range of areas, such as developing new materials with specific properties," says Ellen Moons, Chair of the Nobel Committee for Physics.

John J. Hopfield, born 1933 in Chicago, IL, USA. PhD 1958 from Cornell University, Ithaca, NY, USA. Professor at Princeton University, NJ, USA.

Geoffrey Hinton, born 1947 in London, UK. PhD 1978 from The University of Edinburgh, UK. Professor at University of Toronto, Canada.



https://pni.princeton.edu/people/john-j-hopfield/videos

Hoffmans’ model of Schizophrenia (1989): the pruning hypothesis

❖ Synapses progressively reduce throughout childhood and adolescence.

❖ Schizophrenia: pruning process gone too far.

VOL. 15, NO. 3, 1989 477

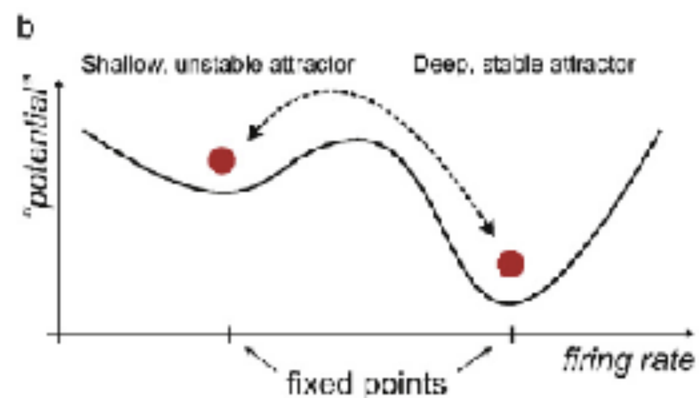
Cortical Pruning and the Development of Schizophrenia: A Computer Model

by Ralph E. Hoffman and Steven K. Dobscha

Abstract

Schizophrenic patients tend to demonstrate reduced cerebral metabolism in frontal areas. Studies of human brain development reveal that synapses in the cerebral cortex are progressively reduced throughout childhood and adolescence, with parallel reductions in cerebral metabolism. This relationship is not surprising, since synaptic density is a primary factor determining regional metabolic requirements. The elimination of

ter being physical phenomena. Computer simulations of neural network interactions are now offering clues about how biologically characterized events can accomplish cognitive tasks. Variants of these models may reveal how disordered neural processes can yield particular kinds of disordered cognition. This report critically examines a number of biological studies of schizophrenic patients in the context of recently emerging knowledge of human brain development. A particular neuroanatomical disturbance



Hoffmans' model of Schizophrenia (1989): the pruning hypothesis

- ❖ Synapses progressively reduce throughout childhood and adolescence.

- ❖ Schizophrenia: pruning process gone too far.

- ❖ Study of pruning in Hopfield nets leads to :
merging of attractors (thought disorder?),
spurious attractors (hallucinations?)..

VOL. 15, NO. 3, 1989 477

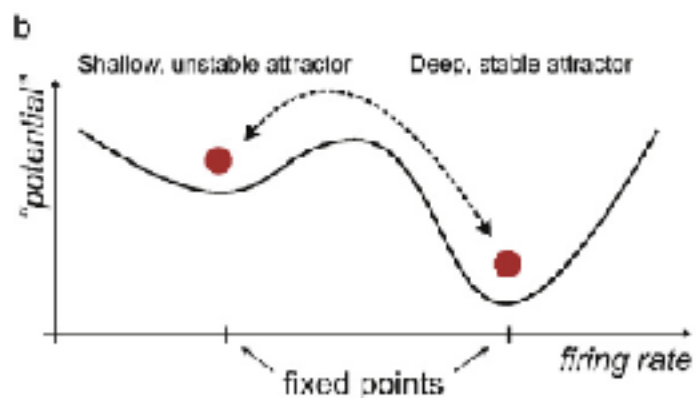
Cortical Pruning and the Development of Schizophrenia: A Computer Model

by Ralph E. Hoffman and Steven K. Dobscha

Abstract

Schizophrenic patients tend to demonstrate reduced cerebral metabolism in frontal areas. Studies of human brain development reveal that synapses in the cerebral cortex are progressively reduced throughout childhood and adolescence, with parallel reductions in cerebral metabolism. This relationship is not surprising, since synaptic density is a primary factor determining regional metabolic requirements. The elimination of

ter being physical phenomena. Computer simulations of neural network interactions are now offering clues about how biologically characterized events can accomplish cognitive tasks. Variants of these models may reveal how disordered neural processes can yield particular kinds of disordered cognition. This report critically examines a number of biological studies of schizophrenic patients in the context of recently emerging knowledge of human brain development. A particular neuroanatomical disturbance



Intermediate Summary

- Schizophrenia is a very serious illness characterised by “positive” (hallucinations, delusions) and negative symptoms.
- One neurobiological correlate of the illness is impairment in working memory.

- Short-term/ working memory: Dynamic process - “Sustained” a.k.a. “Delay” or “Persistent” Activity.

- Attractor Networks as (main) model of working memory / sustained activity

- Hopfield Network as example of a point attractor model (Lab 2).

- Next time: How do we go from here to providing biologically plausible spiking models, comparable to recordings in Prefrontal cortex?