Integrated Information Theory (IIT)

- **Starts from phenomenology**, not from behavioral or neural correlates
- Identifies the essential properties of every experience (**axioms**)
- Derives the requirements that physical systems must satisfy to account for them (**postulates**)
- Has **explanatory**, **predictive**, and **inferential** power

Tononi 2004, 2008, 2012, 2014, 2015
Balduzzi and Tononi, 2008, 2009
Oizumi, Albantakis and Tononi, 2014
Tononi and Koch, 2015
Tononi, Boly, Massimini, Koch, 2016

Integrated Information Theory (IIT)

Phenomenology: axioms

Transcendental properties
of consciousness
(true of every experience)

Physics: postulates

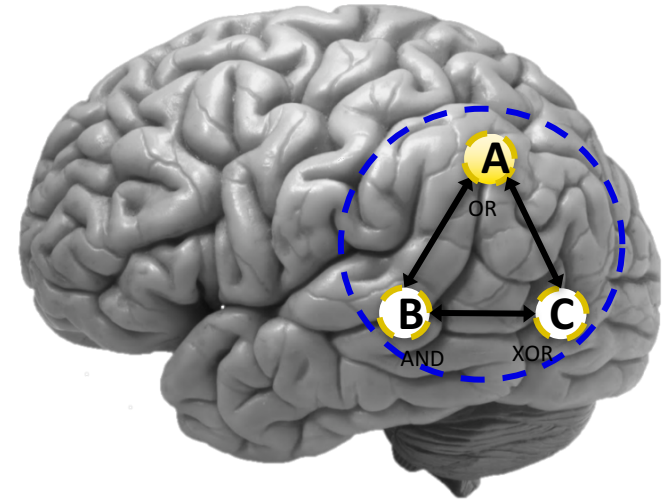
Required properties
of its physical substrate
(necessary and sufficient)

Phenomenal existence

Experience

Physical existence

Cause-effect power



ON *MANIPULATE*

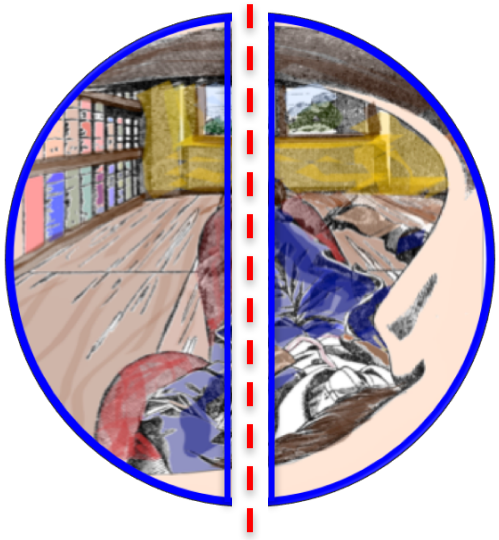
OFF *OBSERVE*

Intrinsicality

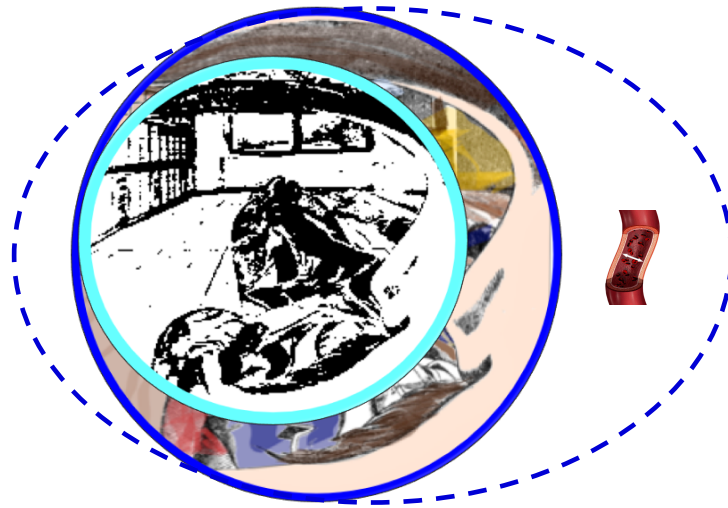
Composition

Information

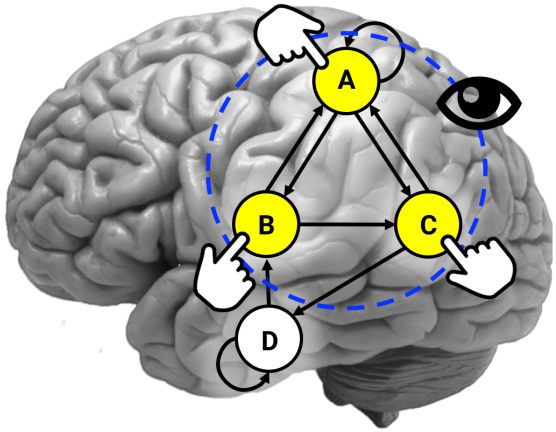
Integration



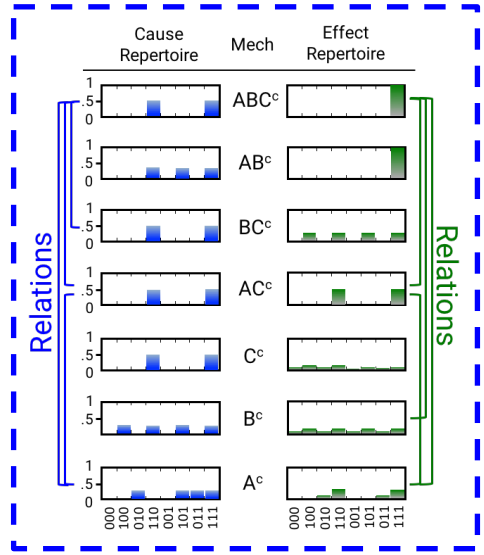
Exclusion



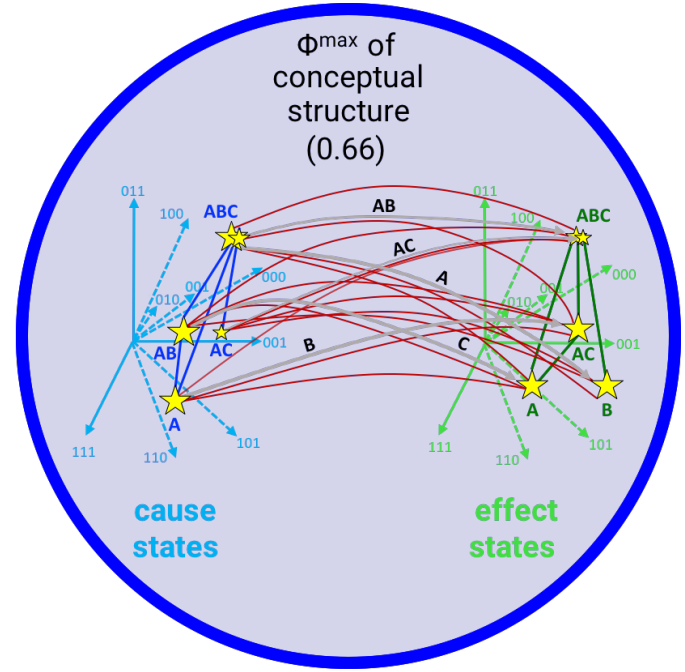
Intrinsicality



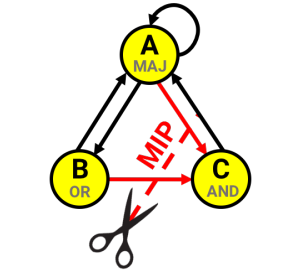
Composition

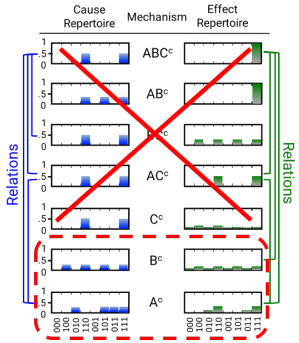


Information

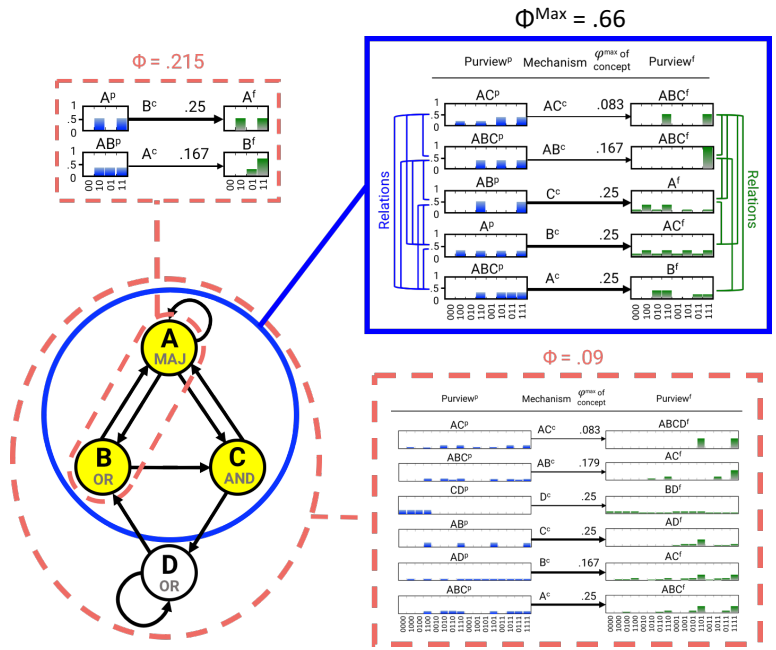


Integration





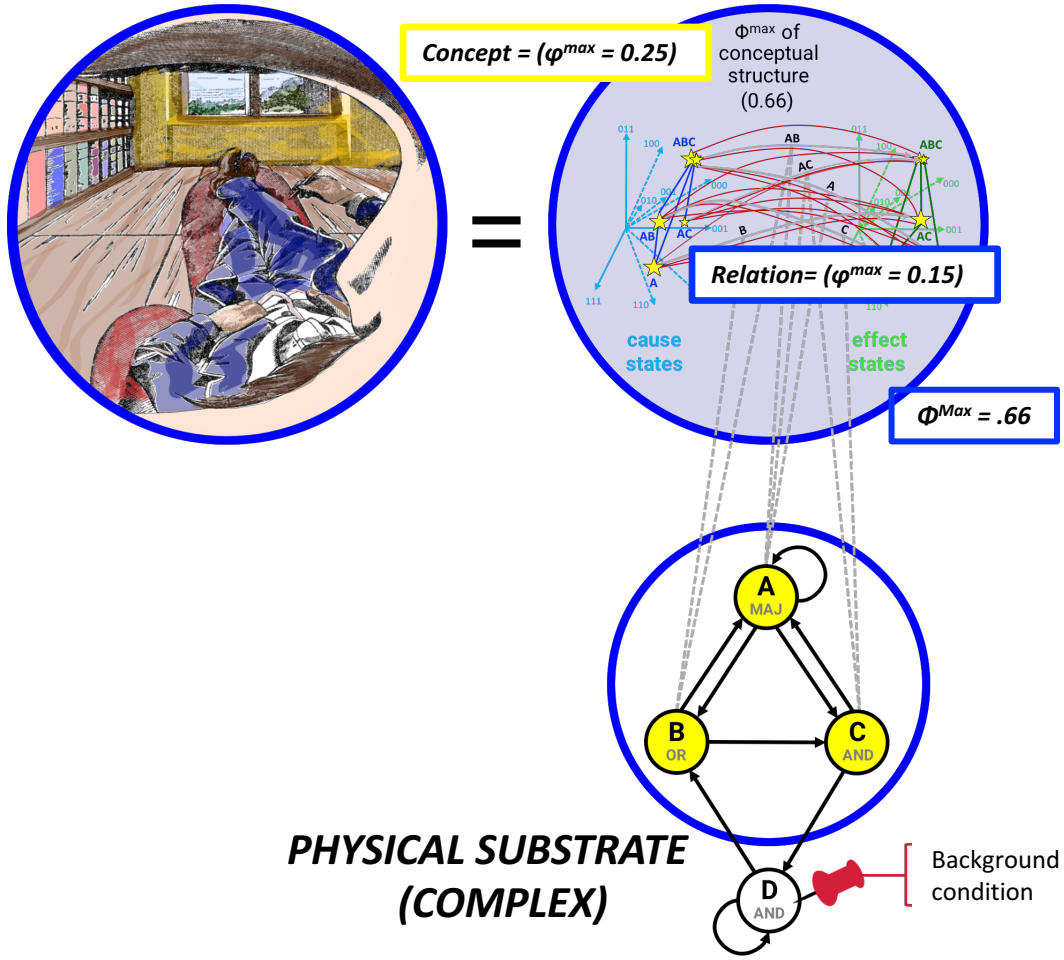
Exclusion



An experience is identical to a conceptual structure - a global maximum of irreducible, intrinsic cause-effect power

specified by a complex of elements in a state

EXPERIENCE ***CONCEPTUAL STRUCTURE***



PHYSICAL SUBSTRATE (COMPLEX)

Quality:

'form' of the conceptual structure

Quantity:

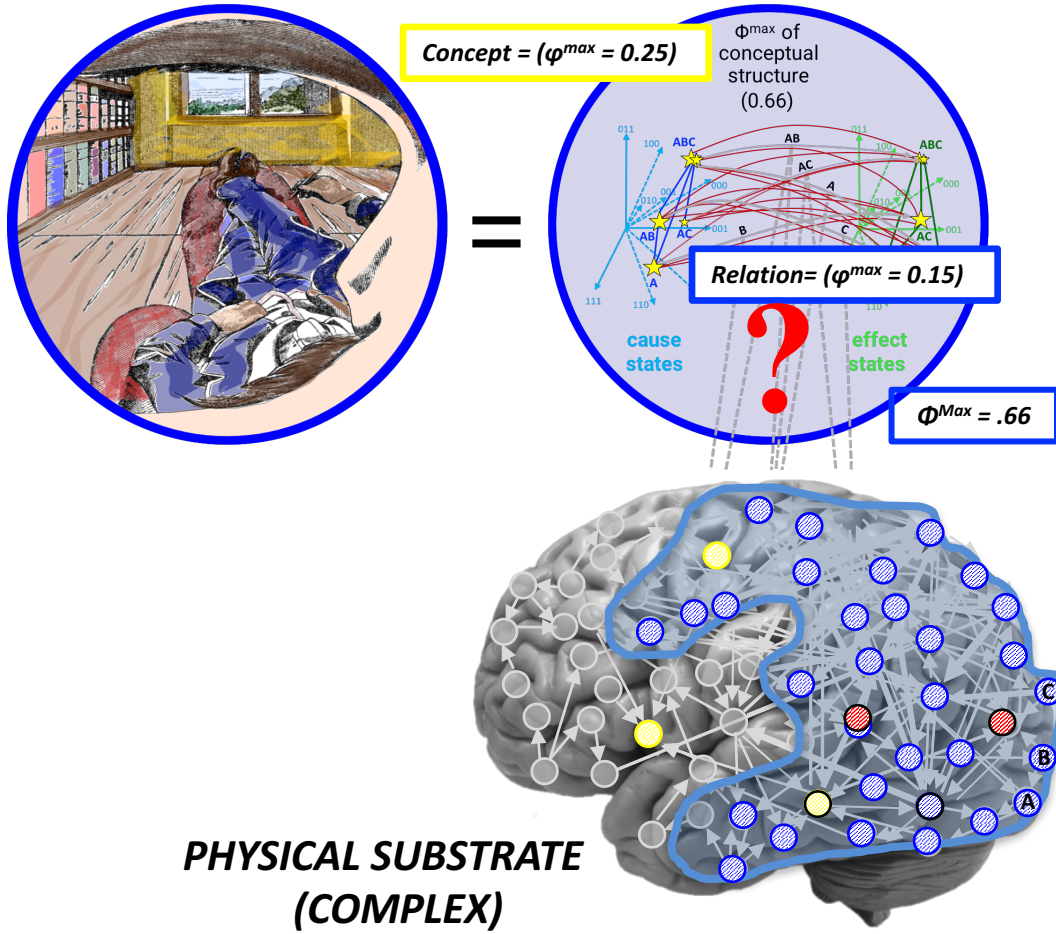
irreducibility (Φ^{max}) of the conceptual structure

Tononi et al
Nature Rev Neurosci 2016

An experience is identical to a conceptual structure - a global maximum of irreducible, intrinsic cause-effect power

specified by a complex of elements in a state

EXPERIENCE ***CONCEPTUAL STRUCTURE***



Quality:

'form' of the conceptual structure

Quantity:

irreducibility (Φ^{max}) of the conceptual structure

Tononi et al
Nature Rev Neurosci 2016

Integrated Information Theory (IIT)

- **Starts from phenomenology**, not from behavioral or neural correlates
- Identifies the essential properties of every experience (**axioms**)
- Derives the requirements that physical systems must satisfy to account for them (**postulates**)
- Has **explanatory**, **predictive**, and **inferential** power

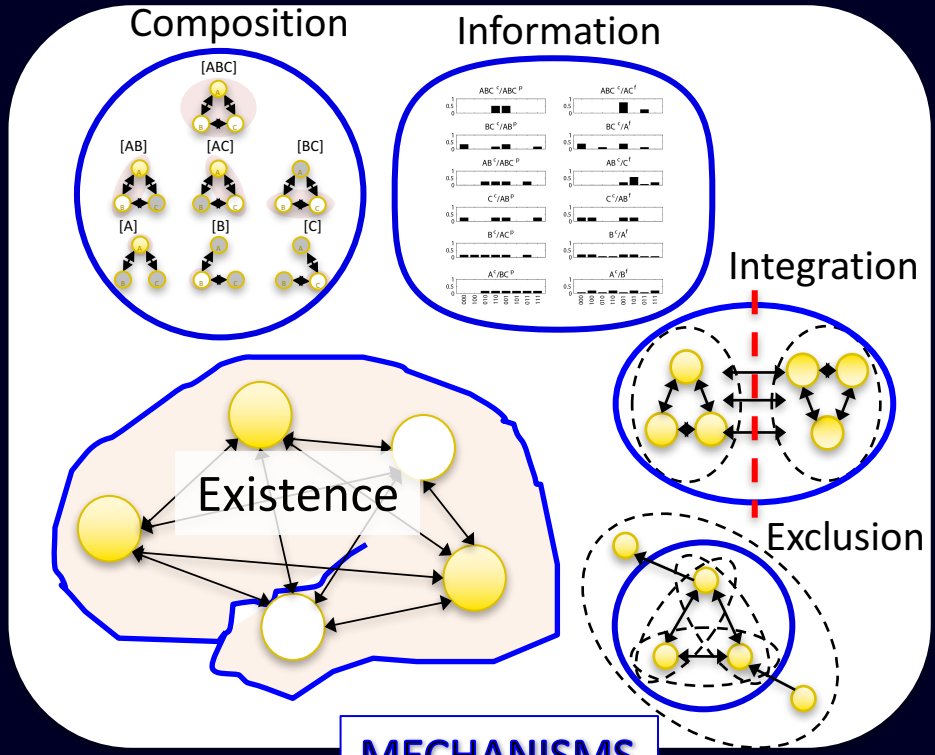
Tononi 2004, 2008, 2012, 2014, 2015
Balduzzi and Tononi, 2008, 2009
Oizumi, Albantakis and Tononi, 2014
Tononi and Koch, 2015
Tononi, Boly, Massimini, Koch, 2016

From phenomenology to mechanisms, and back

PHENOMENOLOGY

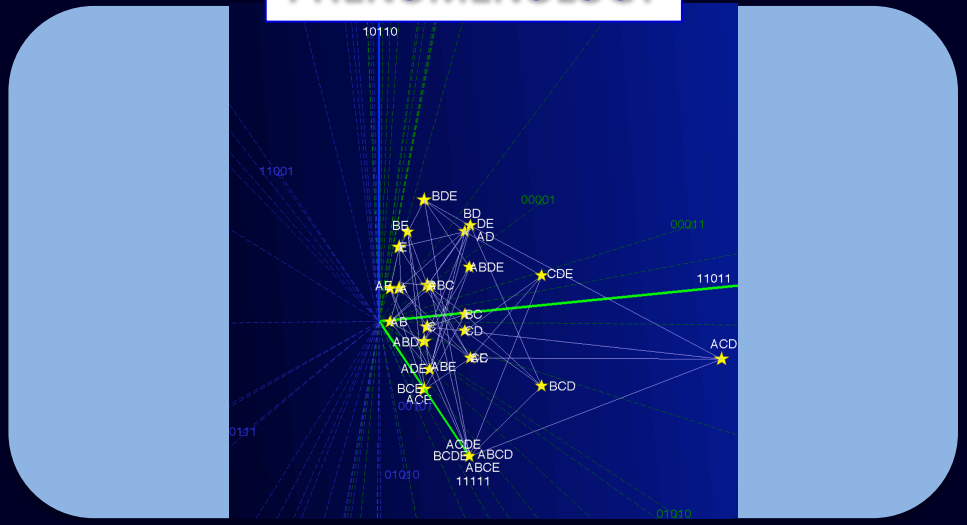
empirical

prediction



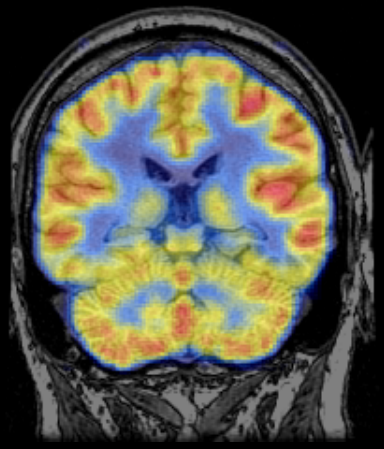
MECHANISMS

inference

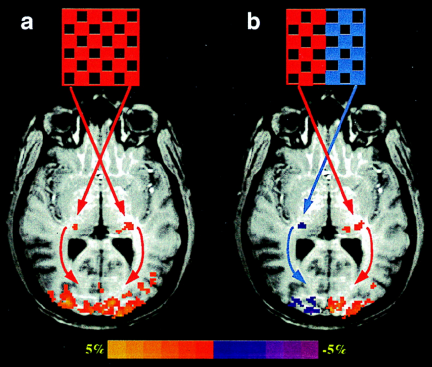


Explanatory power

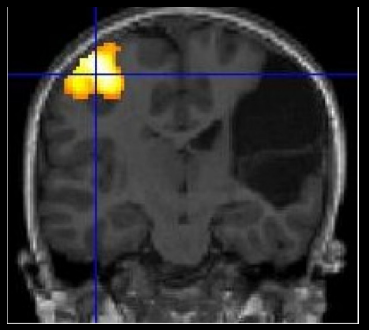
Why not the cerebellum?



Why not afferent pathways?

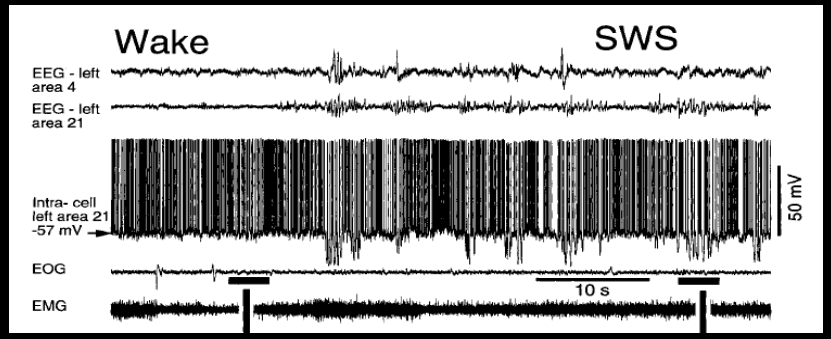


Why not efferent pathways?

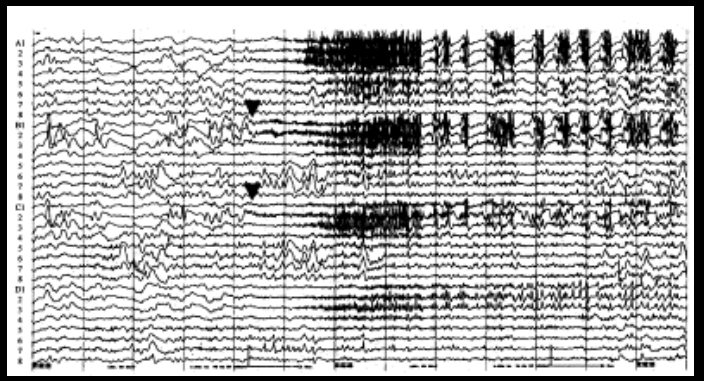


Why not cortico-subcortico-cortical loops?

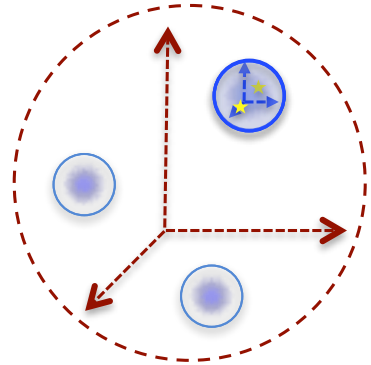
Why not the cortex during deep sleep?

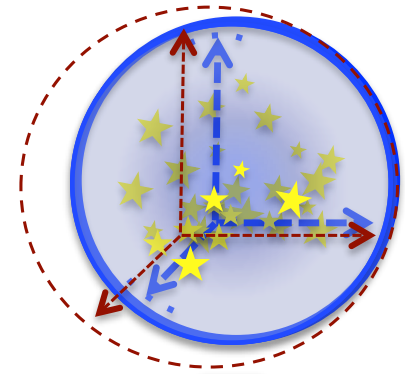


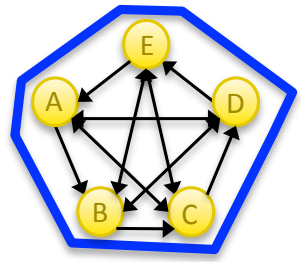
Why not the cortex during a seizure?



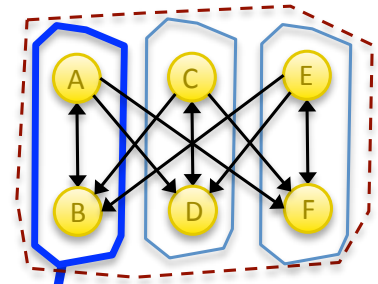
Explanatory power







$\Phi^{Max}_{[AC]-[BDE]} = 10.75$

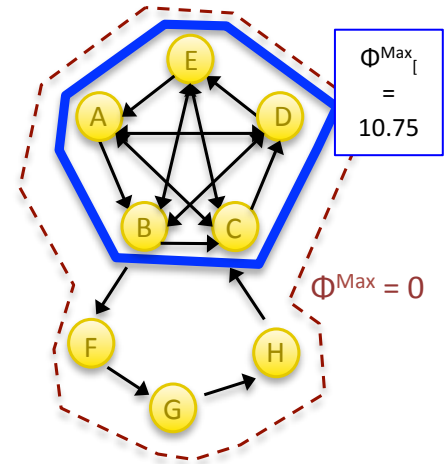


COPYs

ANDs

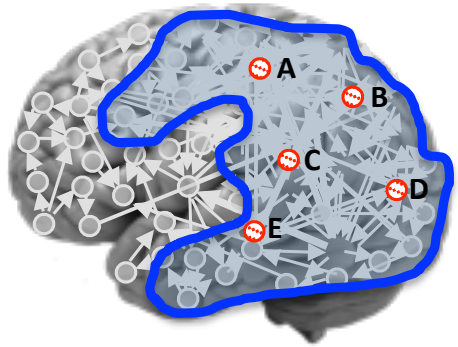
$\Phi_{[AB]-[CDEF]} = 0.53$

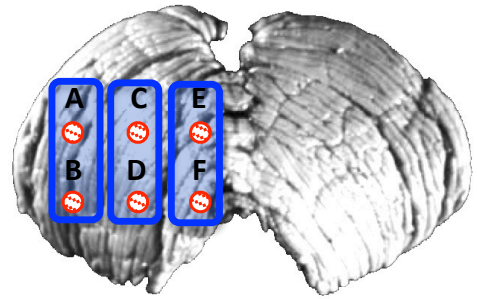
$\Phi^{Max}_{[A\ -[B]} = 1.00$

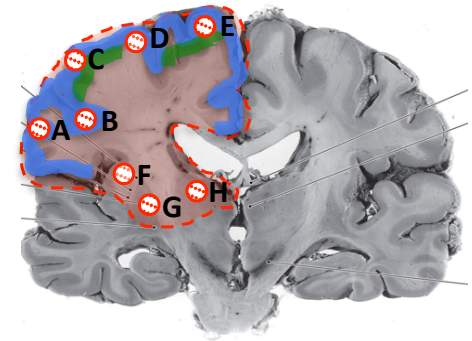


$\Phi^{Max}_{[} = 10.75$

$\Phi^{Max} = 0$

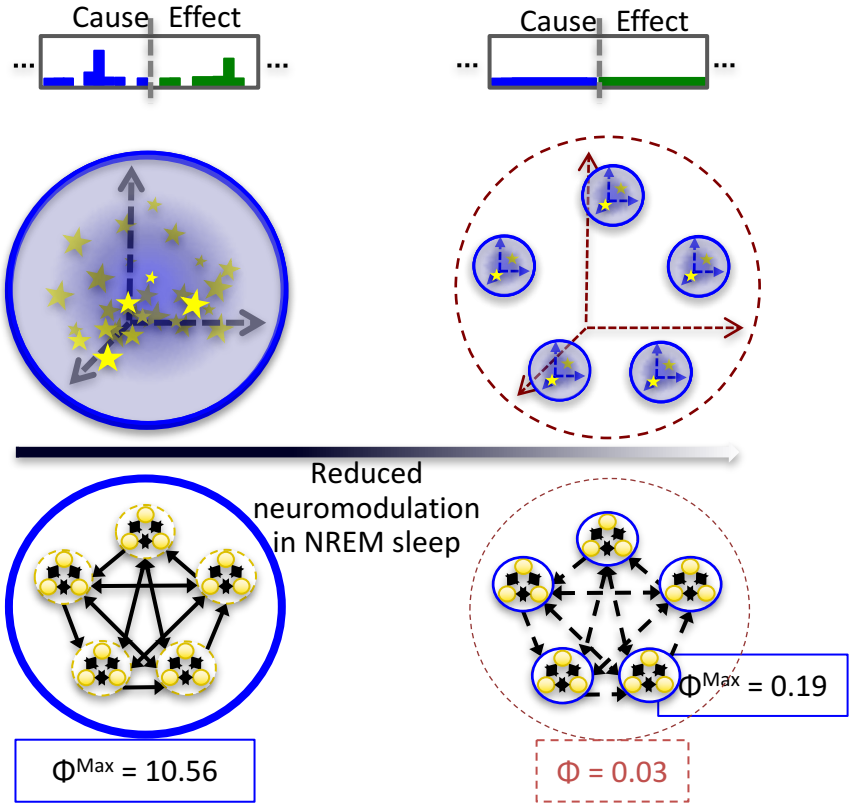






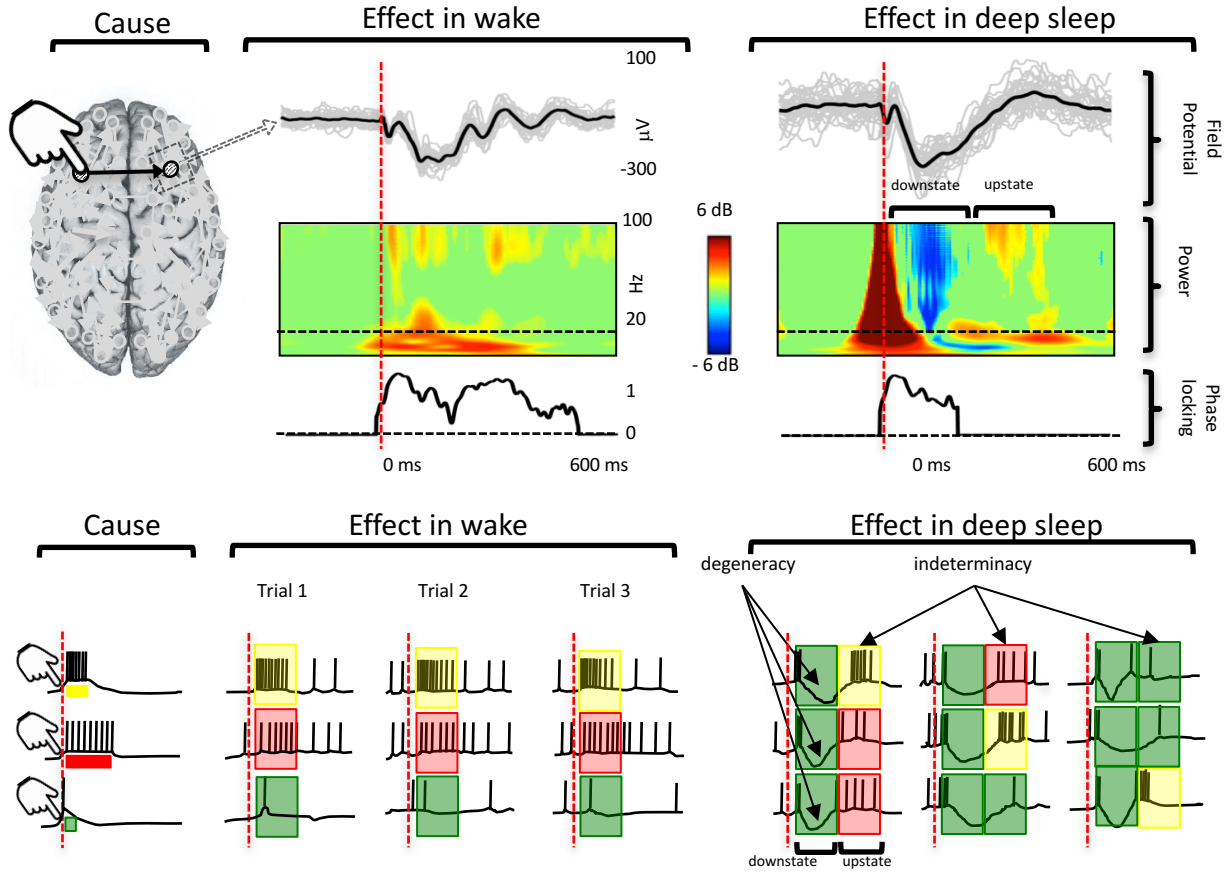
Cerebral Cortex ***Cerebellum*** ***Basal Ganglia loops***

Explanatory power



'Wake

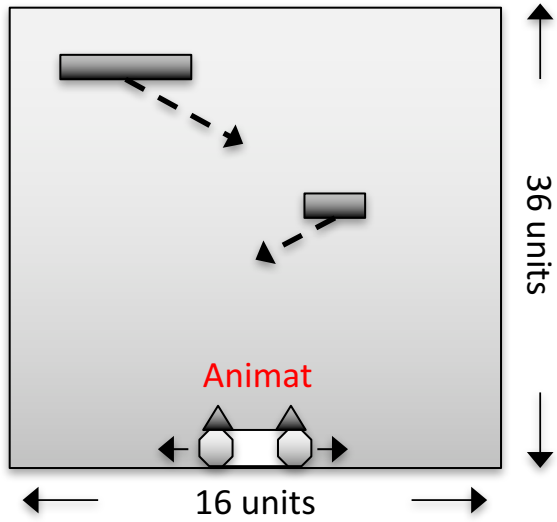
'Deep sleep'

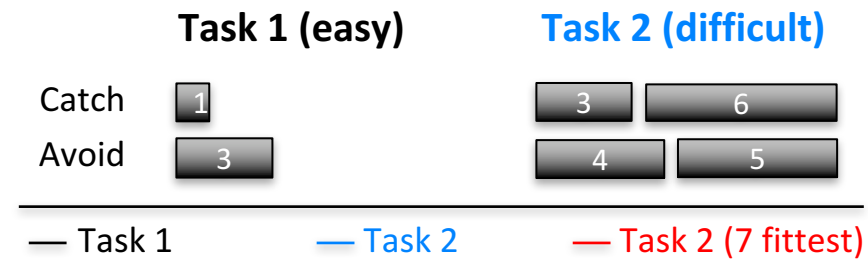


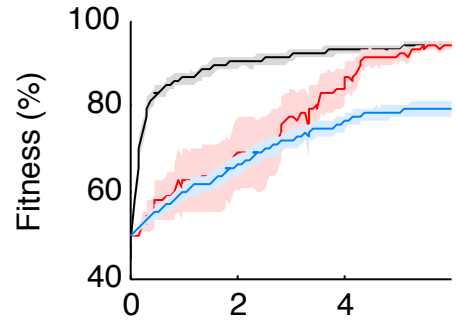
from Pigorini et al., Neuroimage, 2015;
Tononi et al., Nature Rev. Neurosci., 2016

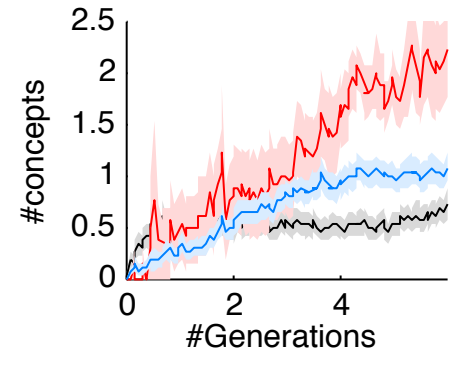

Consciousness as integrated information:
Why did it evolve?

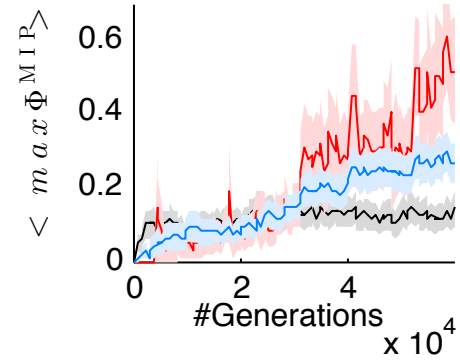
The more difficult the task, the higher the integrated information (<Φ^{max}> and # of concepts) in the fittest animats











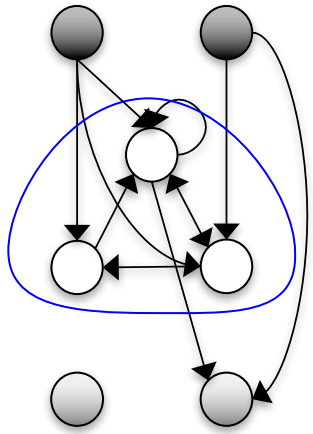
Generation #11264
Fitness: 78.9%

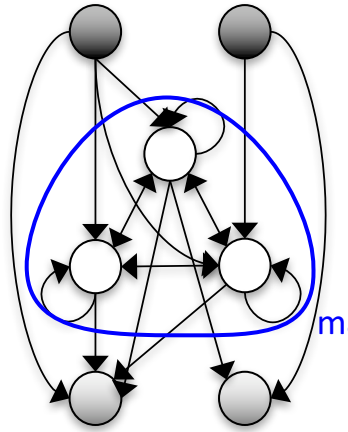
Generation #59904
Fitness: 97.7%

sensors

hidden elements

motors





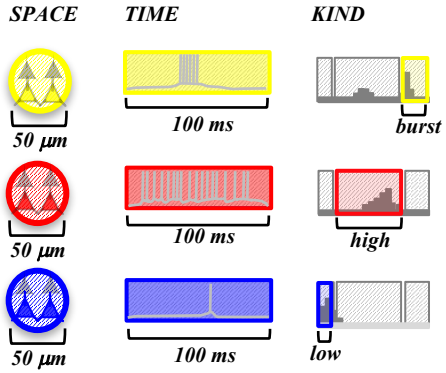
main complex

Main complex:
<Φ^{Max}> = 0.1674
<#concepts> = 2.25

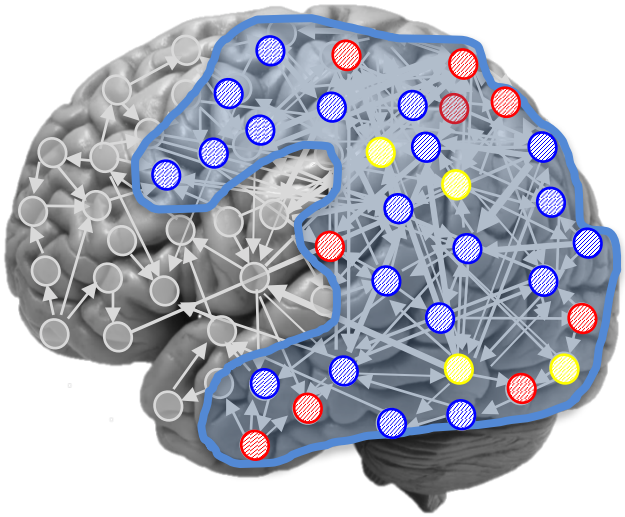
Main complex:
<Φ^{Max}> = 1.13
<#concepts> = 4.4

Albantakis et al., 2014

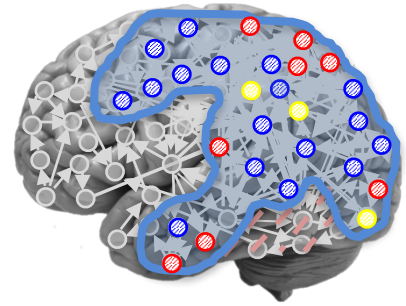
Predictive power



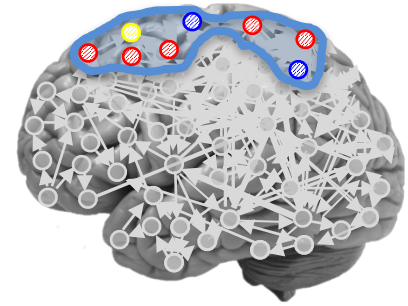
A *The major complex (watching a movie)*



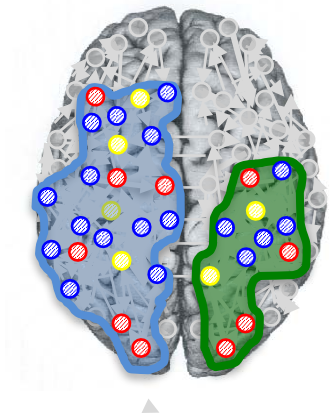
B *It shrinks (absolute achromatopsia)*



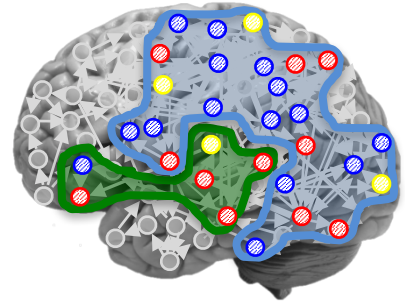
C *It moves (pure thought?)*



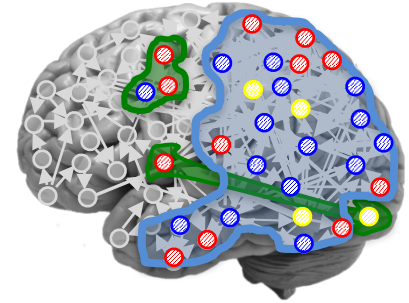
D *It splits anatomically (split brain)*



E *It splits functionally (drive and listen?)*

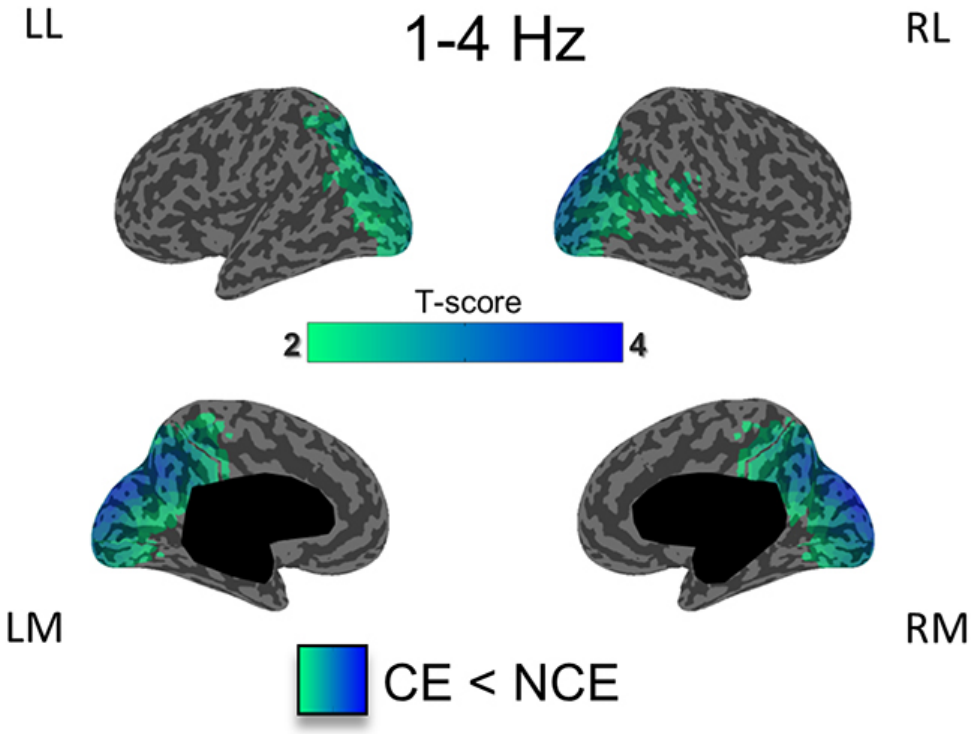


F *It coexists with minor complexes (high-level, "unconscious" feats?)*



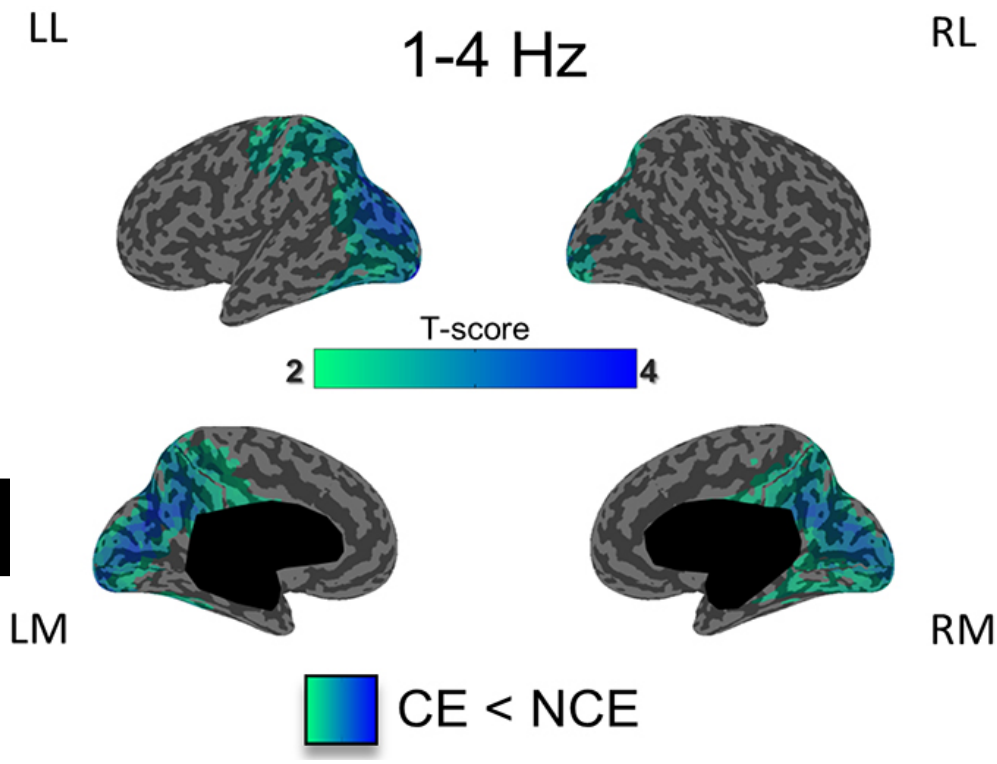
NCC: a within-state, no-task paradigm

NREM sleep



CE: awakening with report of Conscious Experience

REM sleep



NCE: awakening with report of No Conscious Experience

Siclari, Baird, Perogamvros et al., Nature Neuroscience, 2017

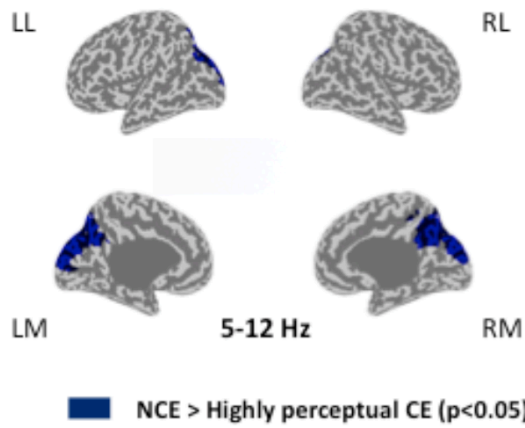
NCC: a within-state, no-task paradigm

Perceptual **Thought- like**

a.

Highly perceptual experiences	Highly thought-like experiences
A documentary about poor children somewhere in Eastern Europe	I was thinking about the possibility of having an erotic dream during this experiment, and that it would be very awkward.
I saw a TV commercial for a car.	I was thinking about the phrase: 'It pays to be wrong sometimes'.
I was watching a medieval play. A man and a woman were lying on the floor, pretending to be dead.	I was thinking about inviting someone to my house.

b.



c.