NETWORK THINKING THEMSELVES

CAMBRIDGE NETWORKS DAY
AUGUST 29, 2019

Danielle S. Bassett

University of Pennsylvania
Department of Bioengineering
Department of Physics & Astronomy
Department of Neurology
Department of Psychiatry
Department of Electrical & Systems Engineering

A story

Roger Deakin: English writer and documentary-maker on water(ways).

"I stared dedicatedly at my shoes, embarrassed that my friend was failing to perform in front of my academic peers. It was only later that I realized it wasn't a failure to perform, but a refusal to conform. Cambridge seminars expect rigor and logic from their speakers: a braced subtlety of exposition and explanation, tested proofs of cause and consequence. But water doesn't do rigor in that sense, and neither did Roger, though his writing was often magnificently precise in its poetry... For Roger, water flowed fast and wildly through culture: it was protean, it was `slip-shape' … and so that was how he followed it, slipshod and shipshape at once, moving from a word here to an idea there, pursuing water's influence, too fast for his notes or audience to keep up with, joining his … watery subjects by means of an invisible network of tunnels and drains."

Waterways of the UK.

Macfarlane, *Landmarks*

I. Knowledge is a network

From Henri Poincare's 1905 *Science and Hypothesis*:

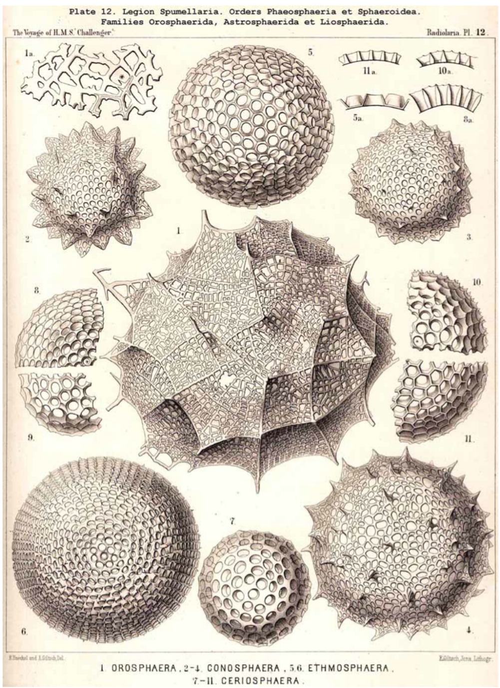
"The aim of science is not things themselves, as the dogmatists in their simplicity imagine, but the relations among things; outside these relations there is no reality knowable."

From Dewey's 1916 *Democracy and Education* (NY: Simon & Brown, 2011):

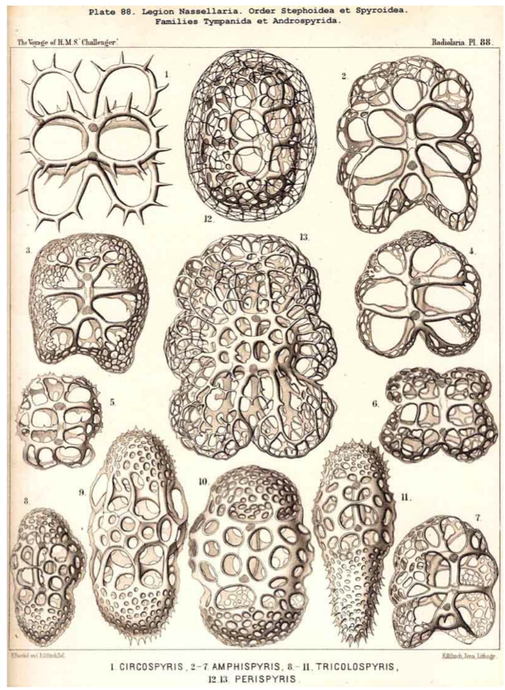
"…[K]nowledge is a perception of those connections of an object which determine its applicability in a given situation. [...] Thus, we get at a new event indirectly instead of immediately - by invention, ingenuity, resourcefulness. An ideally perfect knowledge would represent such a network of interconnections that any past experience would offer a point of advantage from which to get at the problem presented in a new experience" (185).

II. Knowledge is a network learned by example

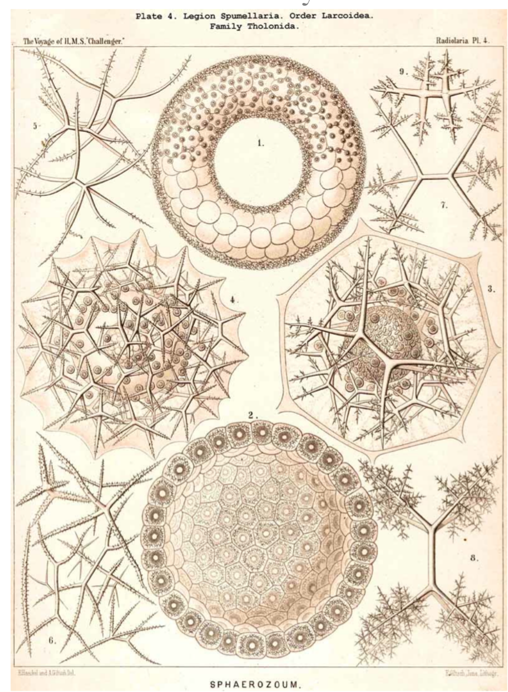
Linear algebra



UK Waterways



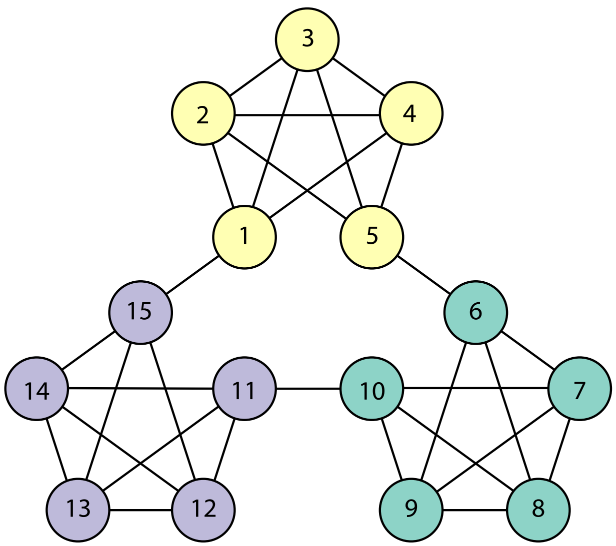
History



Is there an optimal way of walking through a network in lectures, books, papers, etc.?

Ernst Haeckel, *Radiolaria,* microscopic sea organisms

Lectures, Papers, Books: Walks through networks



Let's suppose I have 15 ideas to translate in a class.

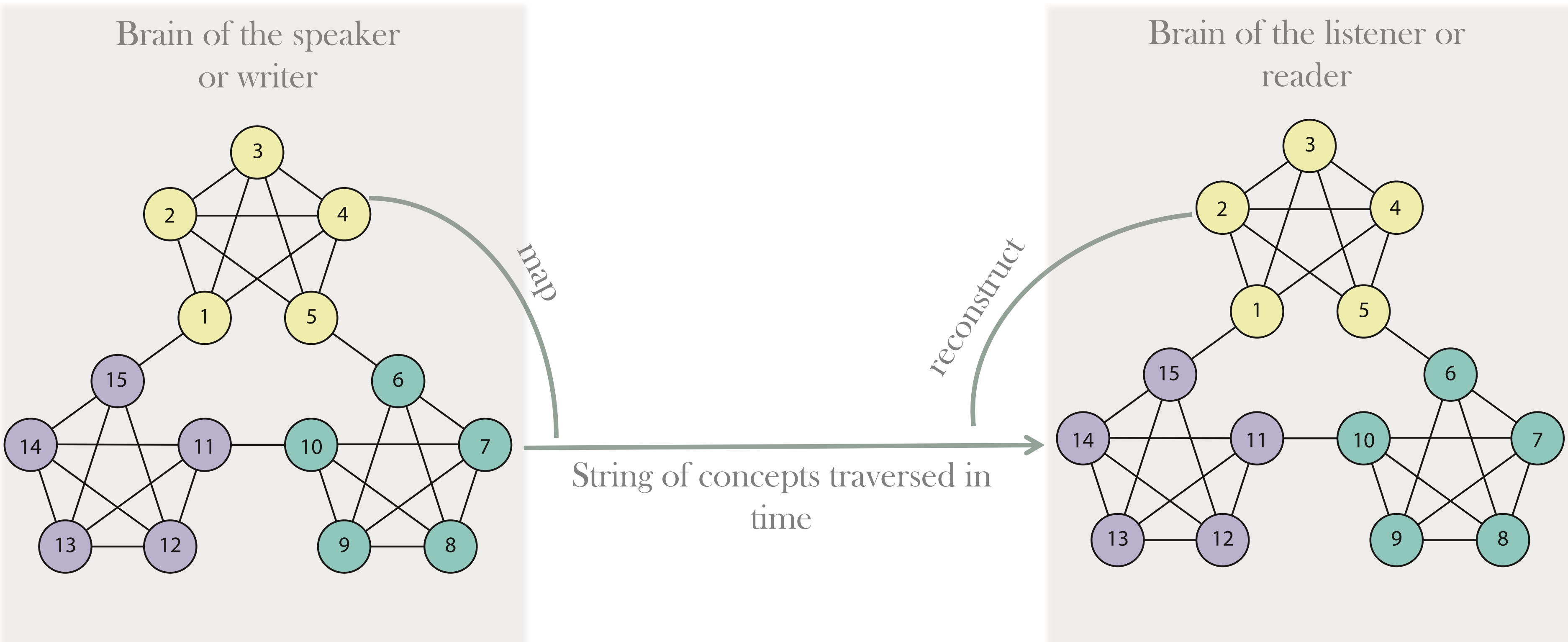
Those 15 ideas are related to one another in a heterogeneous manner, making a network like this ←

But I have to translate that information linearly, because time is one-dimensional and uni-directional.

How should I do it in a way that maximizes learning?

A "good walk" minimizes reconstruction error

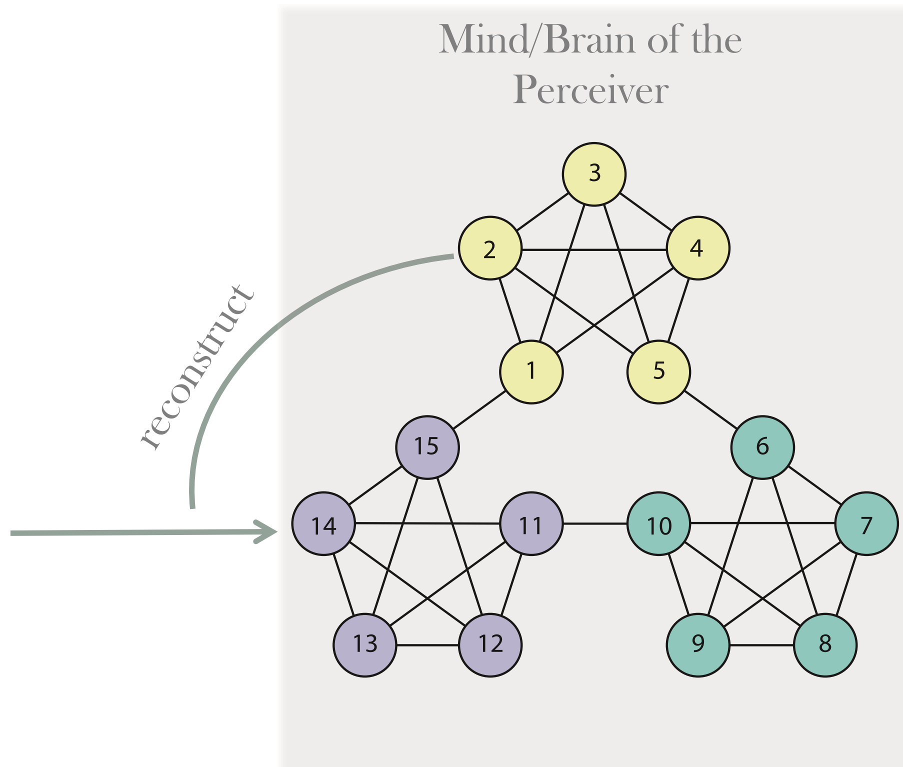
and maximizes perception of the network's topology



One word after another ....
One line after another ...

The problem of inferring the patterns of pairwise dependencies from incoming streams of data allows us to:

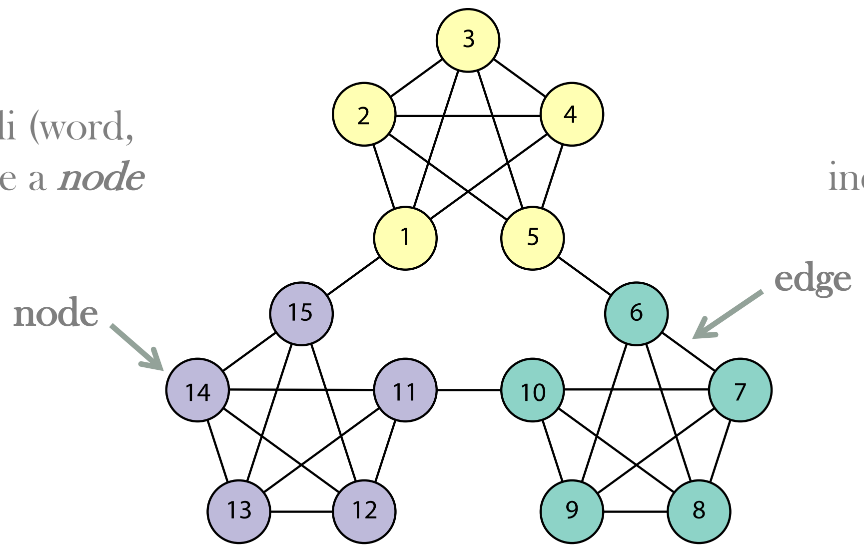
Learn language
Segment visual events
Parse tonal groupings
Parse spatial scenes
Infer social networks
Perceive distinct concepts



Can we measure perception of network topology

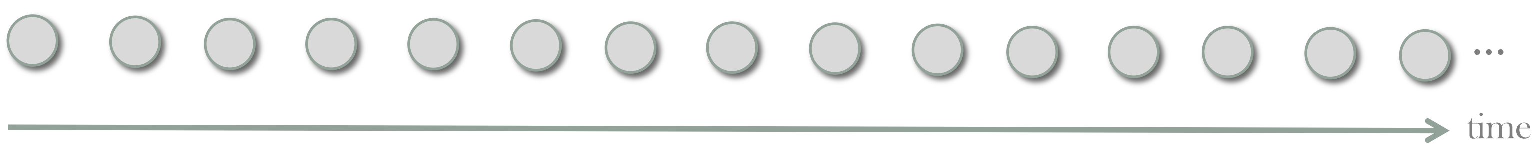
in a continuous stream of stimuli?

Let each specific stimuli (word, image, or movement) be a ***node in a graph***.



Let each ***edge in the graph*** indicate an allowable transition between nodes.

Construct a sequence of stimuli by a random walk on the graph.



At each stimuli, require the participant to perform a task, so that their time-to-react can be used as a measure of how well that edge in the graph was learned.

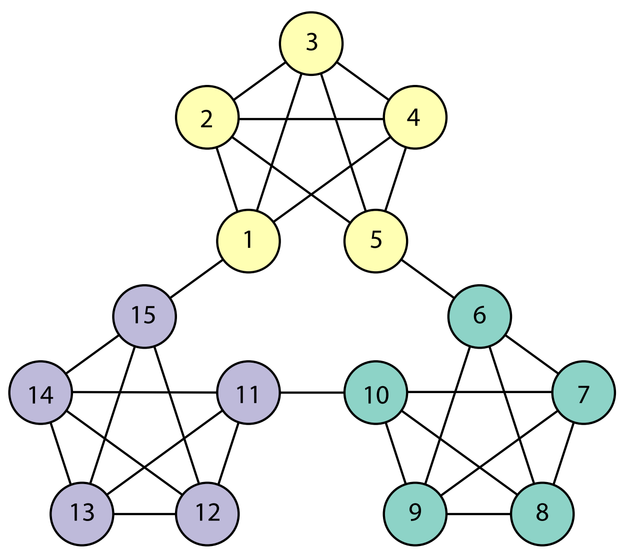
Karuza et al. 2016 *Trends in Cognitive Science*

Danielle S. Bassett

What do we know about this problem?

From work in the field of statistical learning and the study of artificial grammars, we know that humans are sensitive to transition probabilities.

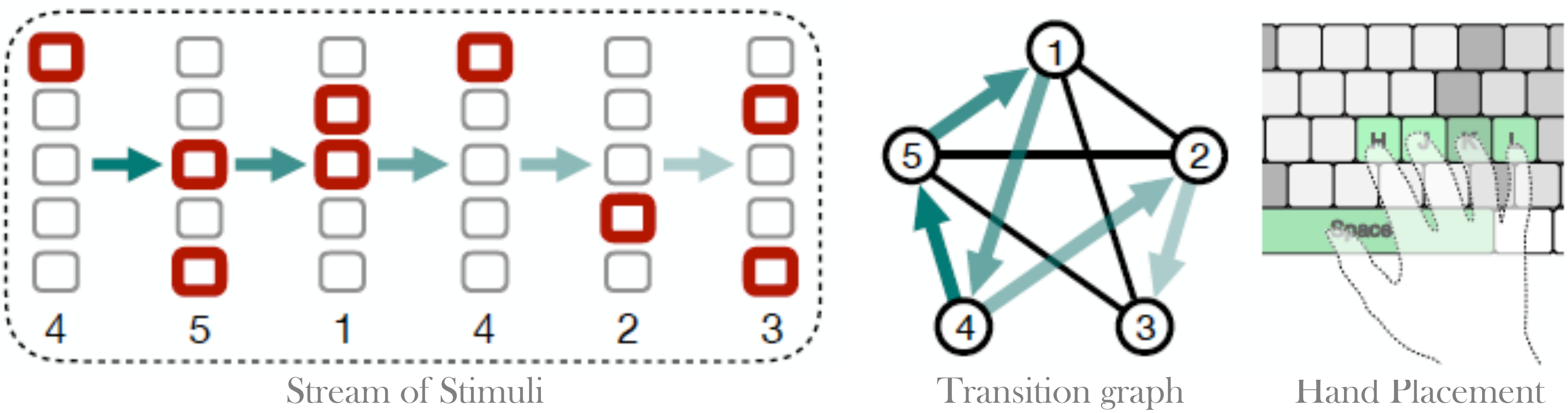
Then what would we predict about the graph below?



Because every edge has a transition probability of 0.25, human expectations should be equivalent across all transitions, and thus so should human reaction times.

Example experimental setup

1. Motor: *Kahn et al. 2018 Nature Human Behavior*

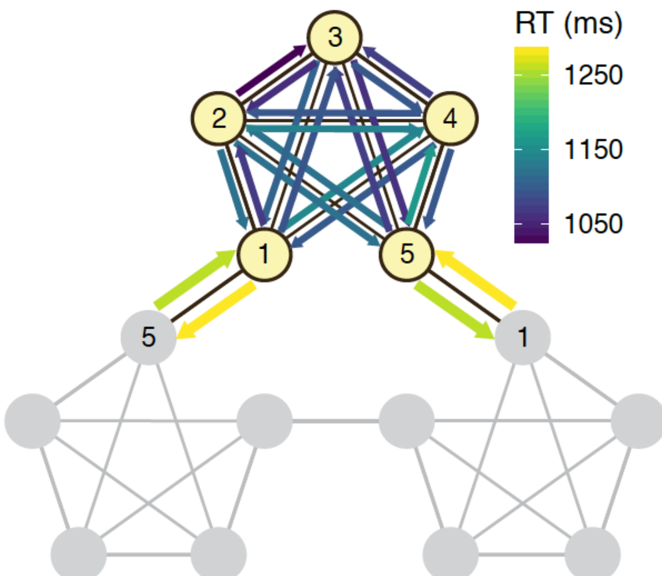


2. Visual: Karuza et al. 2017 *Scientific Reports*
3. Social: Tompson et al. 2018 *Journal of Experimental Psychology; Learning Memory & Cognition*

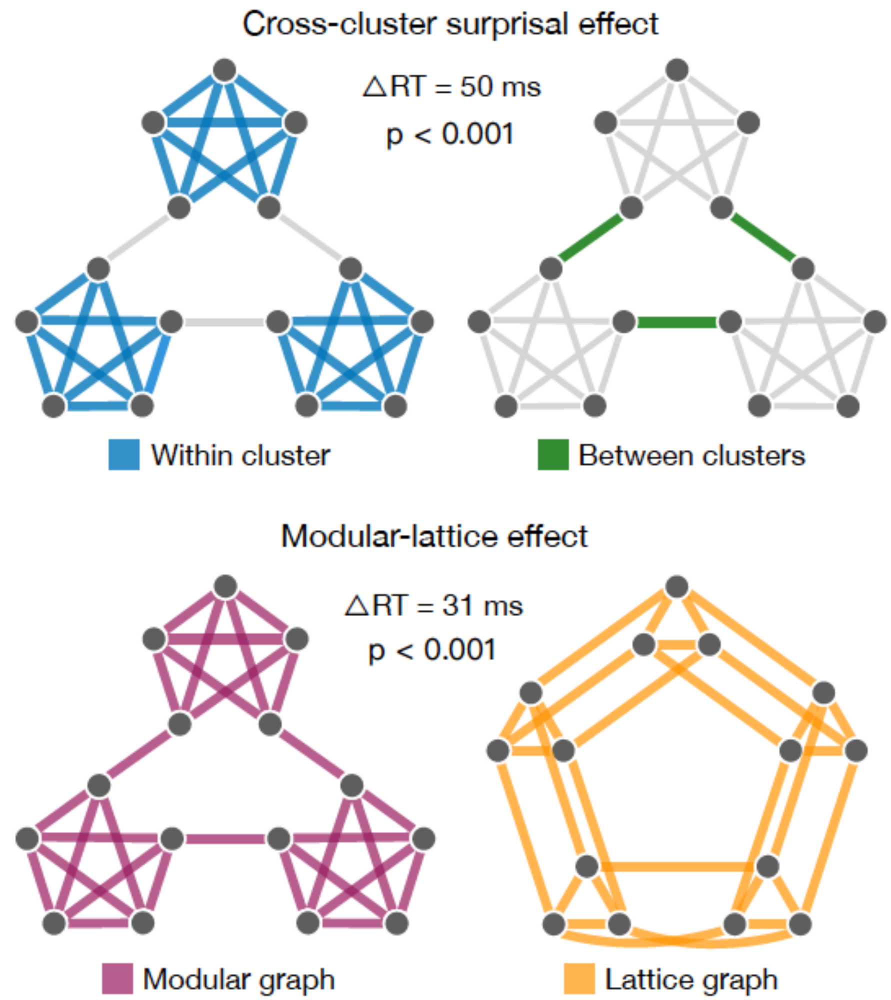
Perception of higher-order network structure

in continuous streams of stimuli

Measure Reaction Time (RT)
Per Transition



Striking slowing at cluster boundaries indicating graph learning



Creating expectations of transitions

We build expectations about a network structure with a counts matrix n_{ij}

Free energy principle: brain minimizes errors & computational complexity.

Probability of recalling $X_{t-\Delta t}$ rather than X_t is $Q(\Delta t)$.

The error of a candidate probability distribution is $E(Q) = \sum_{\Delta t} Q(\Delta t)\Delta t$

Complexity of the error distribution is the entropy $-S(Q) = \sum_{\Delta t} Q(\Delta t) \log Q(\Delta t)$

Total cost of the distribution is its free energy: $F(Q) = \beta E(Q) - S(Q)$

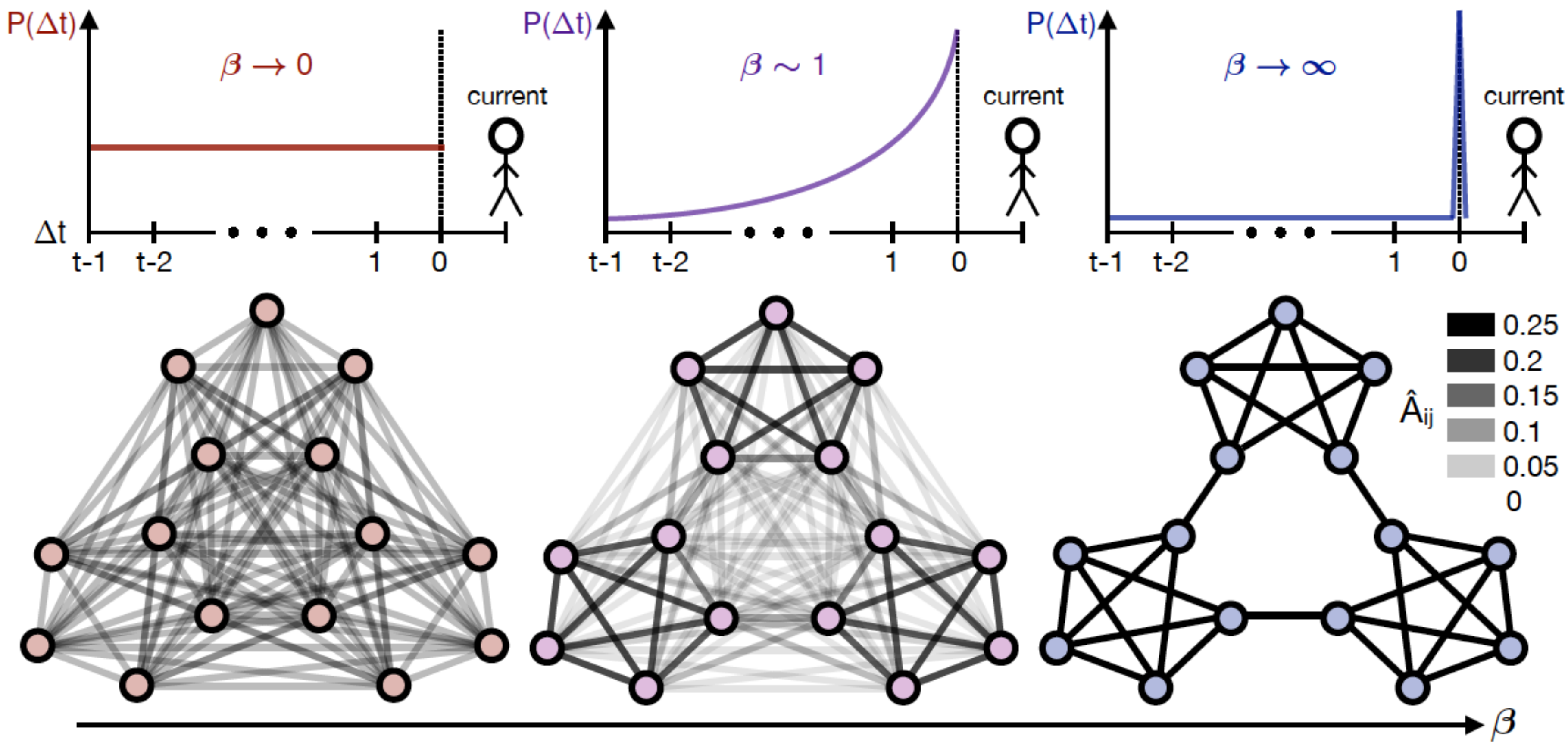
Distribution that minimizes free energy principle is Boltzmann distribution $P(\Delta t) = \frac{1}{Z} e^{-\beta \Delta t}$

From a poor memory arises biases in learning

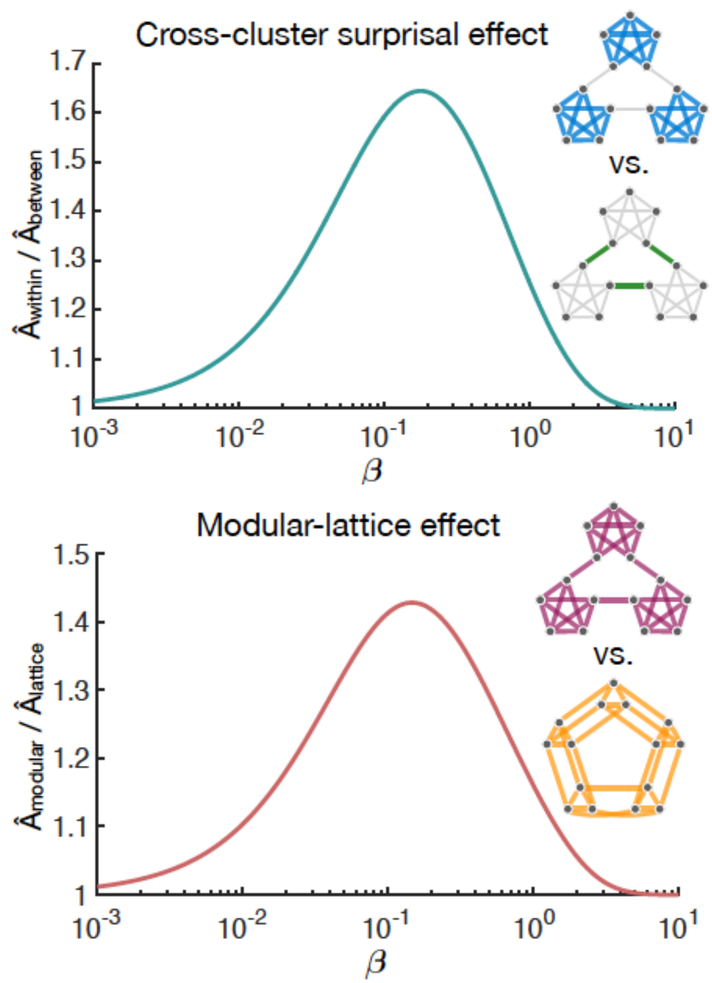
No memory;
Minimizes mental resources

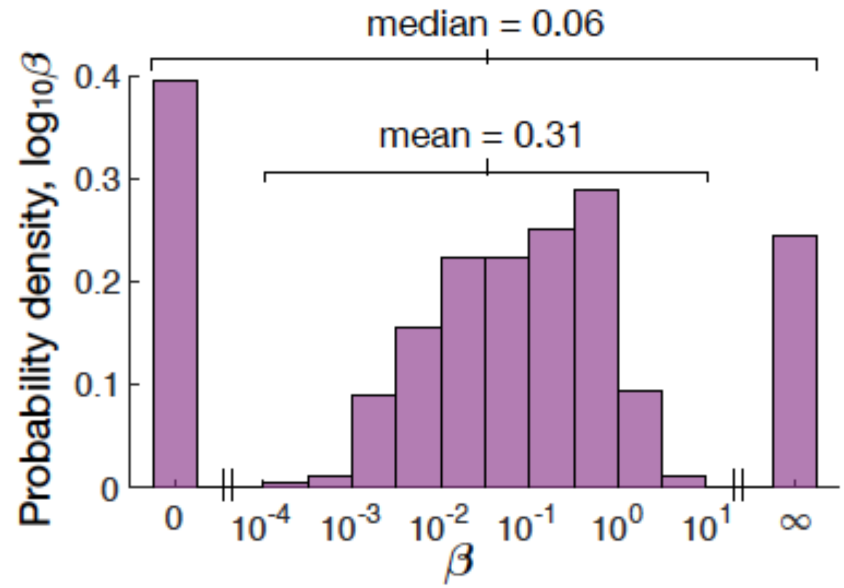
Poor memory

Perfect memory;
Maximizes mental resources



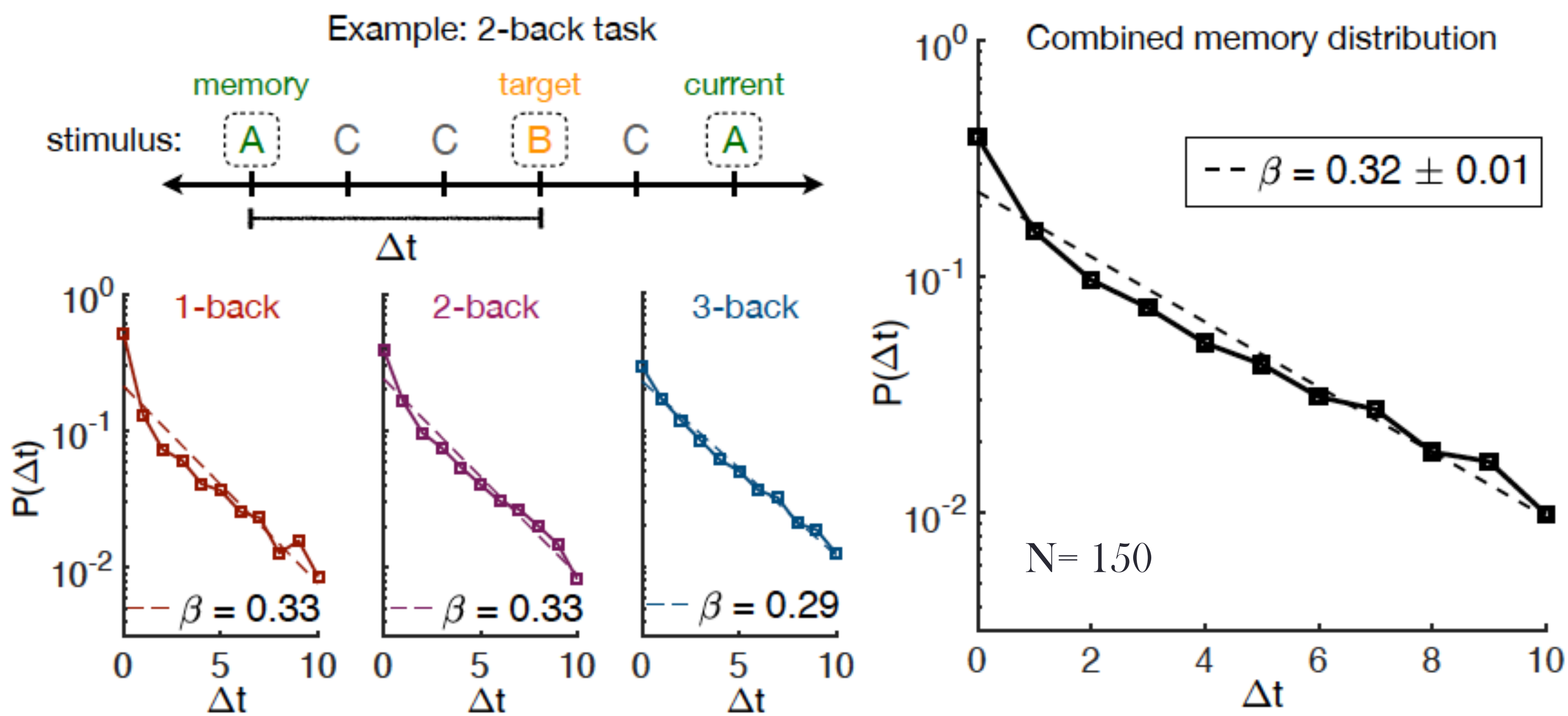
Measuring inverse temperature from RT data





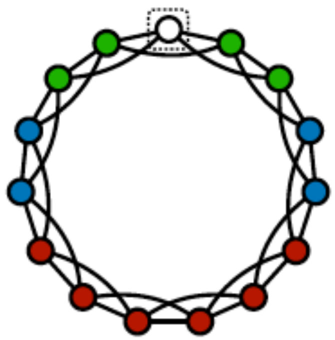
N= 71; 243; 44

Measuring the memory distribution from n-back



The effects of network violations

• Ring graph:



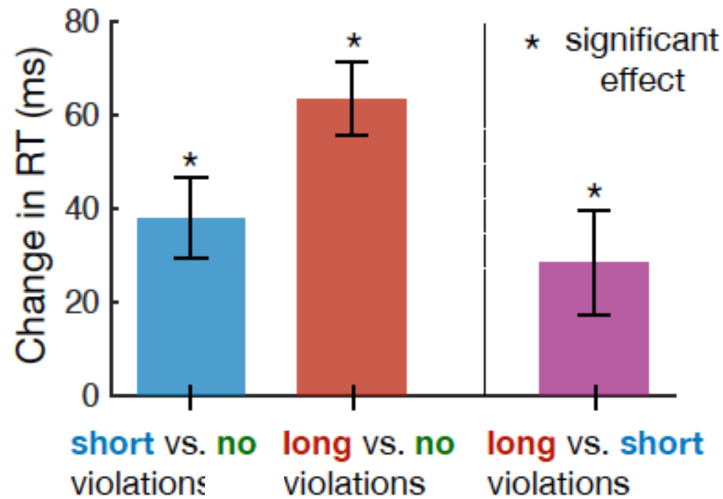
Topological distance:

0 (current node) 2 (short violation)

1 (no violation) 3,4 (long violation)

N= 99

• Empirical results:



Humans are more surprised by stimuli from farther away on the ring than closer, indicating their implicit perception of the network topology.

Searching for design rules

What is the optimally learnable graph? Does it have a topology that is common in language or nature? (Lynn et al. In Preparation) Or in well-written papers? (Chai et al. arXiv:1810.10534) Or in well-written textbooks? (Christianson et al. arXiv)

Do different humans prefer to learn information on different graph architectures? (Lynn et al. 2018, arXiv:1805.12491)

When we employ the processes of graph learning to grow our knowledge networks, do we ever form gaps in knowledge, and if so why? And what do we do with them? (Sizemore et al. 2018 *Nature Human Behavior*) Do different styles of gap-y learning relate to different styles of curiosity? (Lydon-Staley 2019 Psyarxiv)

Dr. Lizz Karuza,
Now Asst. Prof of
Psychology at PSU

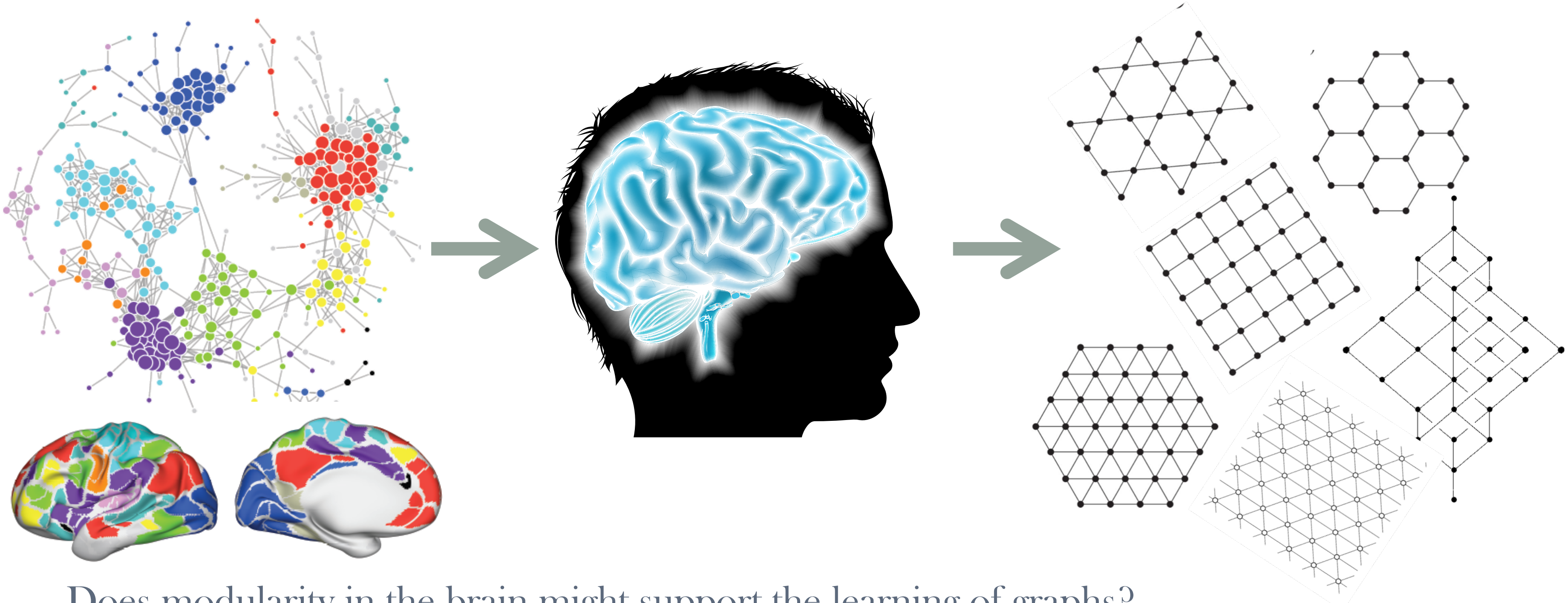
Ari Kahn,
Graduate
Student in
Neuroscience

Chris Lynn,
Graduate
Student in
Physics

Karuza et al. 2016 *Trends in Cognitive Science*

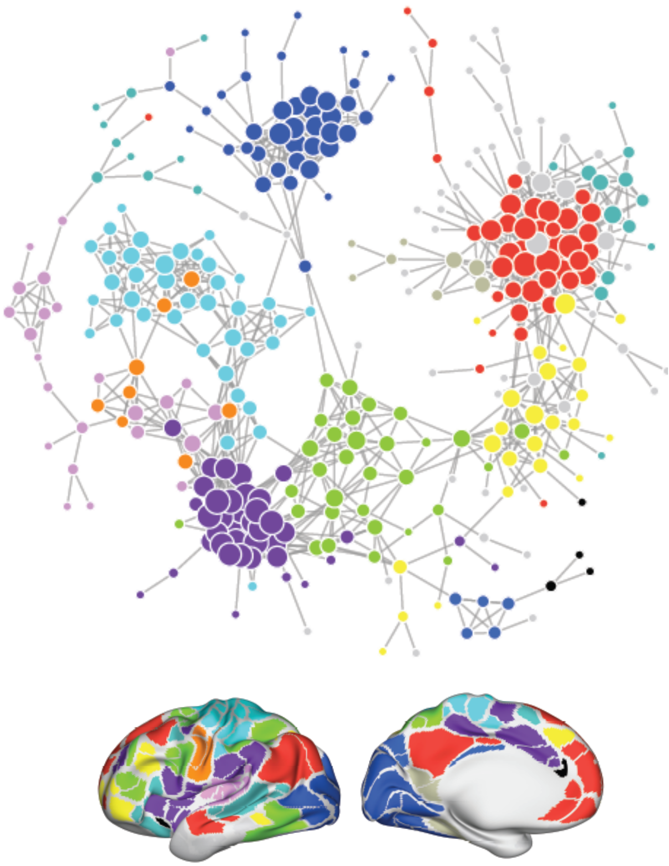
Brain network processes supporting learning

"Mind thinks itself because it shares the nature of the object of thought; for it becomes an object of thought in coming into contact with and thinking its objects, so that mind and object of thought are the same." Aristotle, Metaphysics, Book XII, 7, 1072 b 20



Does modularity in the brain might support the learning of graphs?
Might differences in modularity explain differences in the ability to learn?

Theoretical & Computational Challenges



Challenge: Parsimoniously representing and describing complex connectivity patterns.

- Network models
  - Bassett, Zurn, Gold (2018) *Nat Rev Neuro*

Challenge: Detecting modular structure in network models of brain connectivity.

- Modularity maximization (NP Hard)
  - Meunier et al. (2009) *NeuroImage*

$$Q = \sum_{ij} [A_{ij} - P_{ij}] \delta(\sigma_i \sigma_j)$$

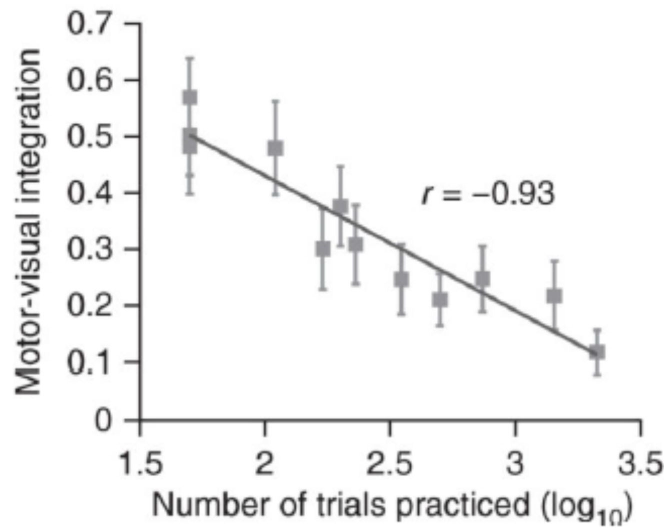
Challenge: Detecting evolving modules.

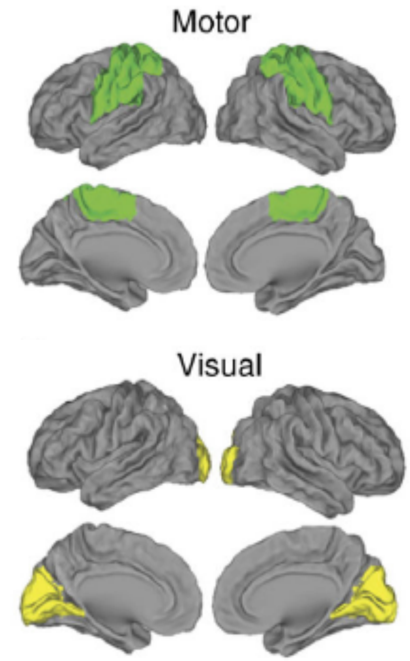
- Multilayer modularity maximization
  - Mucha et al. (2010) *Science*

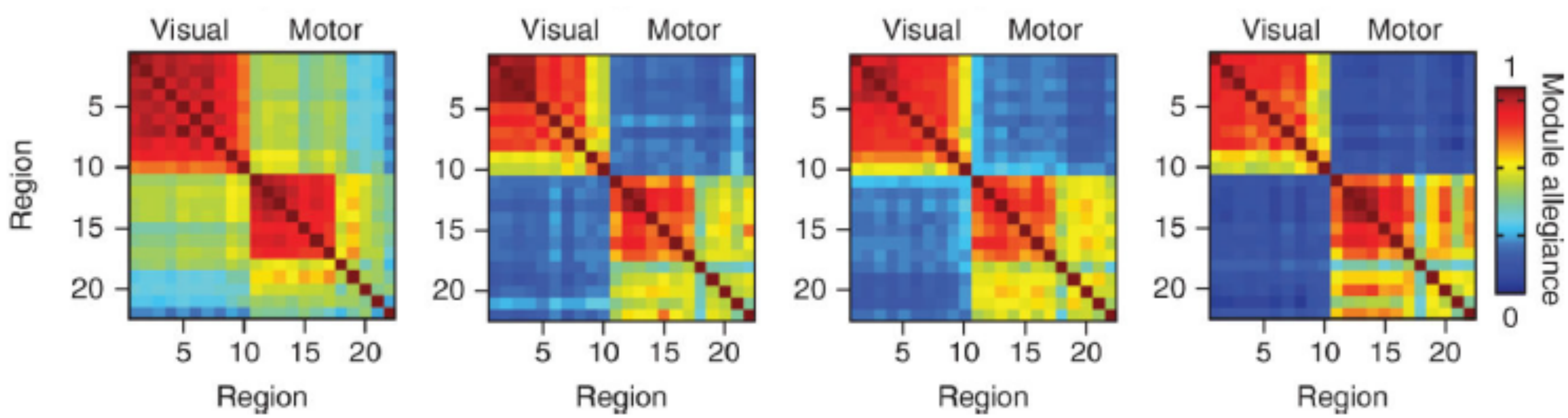
$$Q_{multi} = \frac{1}{2\mu} \sum_{ijls} [(\mathcal{A}_{ijl} - \gamma_l \mathcal{P}_{ijl}) \delta_{lm} + \omega_{jlm} \delta_{ij}] \delta(c_{il}, c_{jm})$$

Module Autonomy in Sequence Learning

The coherence between motor and visual modules decreased markedly with training, suggesting a growing autonomy.

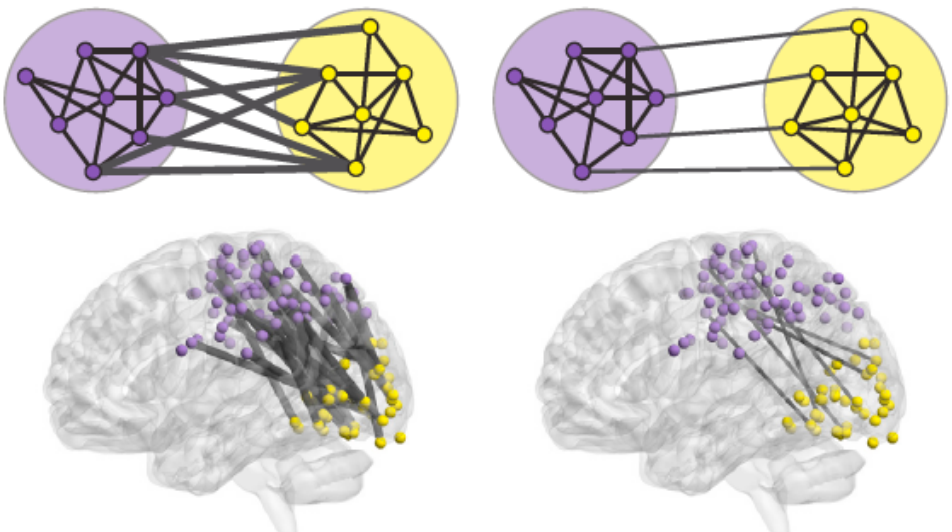






General relevance of modularity for learning

Hypothesis: Networks that can flexibly adapt are those with greater modularity.



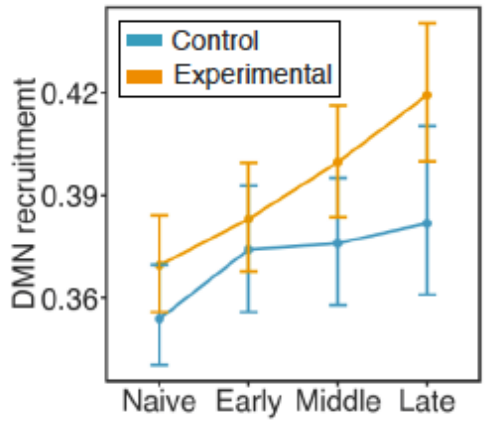
Constrained modules; slow learner

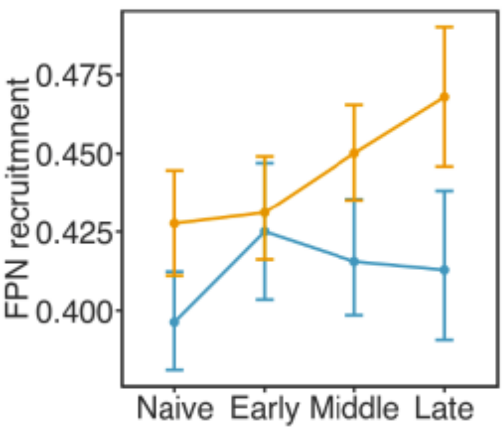
Flexible modules; Fast learner

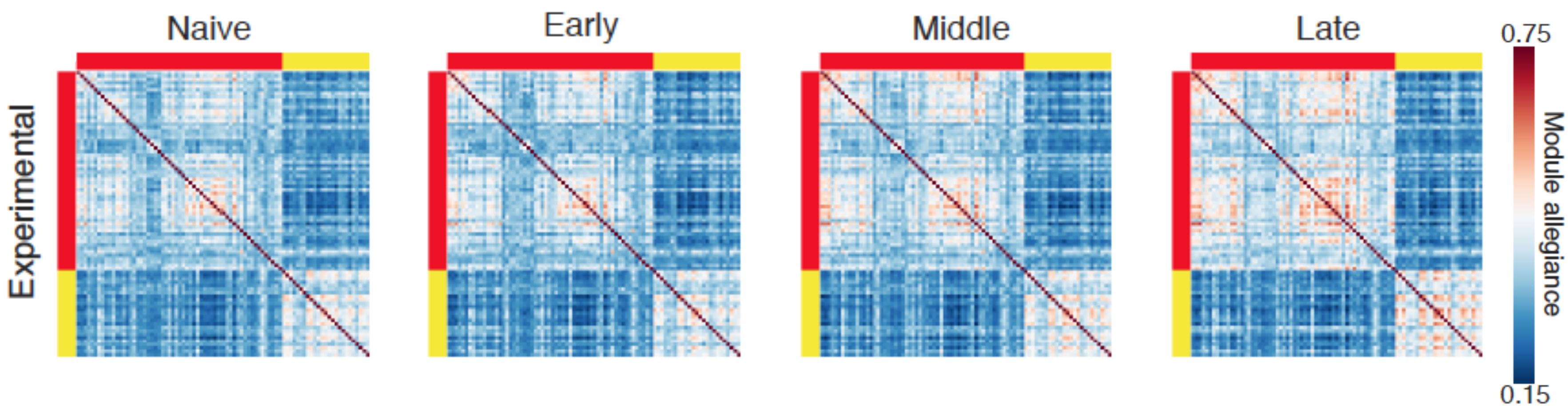
1. Flexible modules support swifter learning over 3 days; *Bassett et al. 2011 PNAS*
2. Swift learning is associated with flexible segregation of modules over 6 weeks; *Bassett et al. 2015 Nature Neuroscience*
3. Segregation of modules at rest predicts learning 6 weeks in the future; *Mattar et al. 2018 NeuroImage*

Module strengthening with dual n-back training

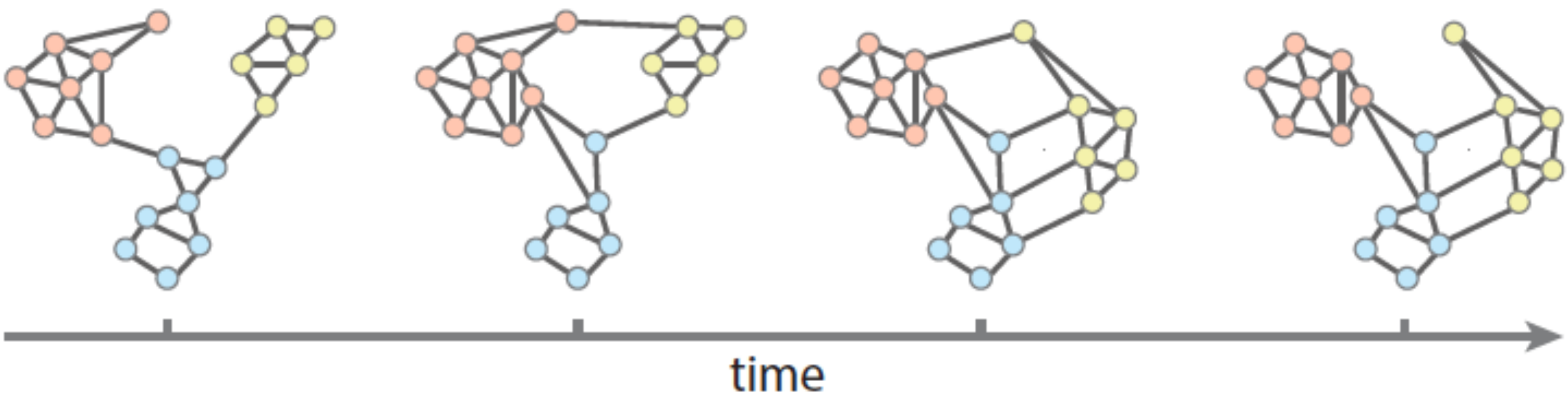
Large-scale collections of brain regions (modules) change in their coherent activity with training, providing coarse-grained markers of function.







Flexible modularity supports learning (& executive function)



Flexibility in network modules is predicts individual differences in:

Visuo-motor learning (Bassett et al. 2011 *PNAS*)
Cognitive flexibility (Braun et al. 2015 *PNAS*)
Working memory (Braun et al. 2015 *PNAS*)
(Shine et al. 2016 *Neuron*)
Learning rate (Gerraty et al., 2018, *J Neurosci*)
Future learning (Mattar et al. 2018 *NeuroImage*)
Planning & reasoning (Pedersen et al. 2018 PNAS)

Medication (Braun et al. 2016, *PNAS*)
Positive mood (Betzel et al. 2017 *Sci Rep*)
Amount of sleep (Pedersen et al. 2018 PNAS)

Searching for design rules

Why do some types of learning induce changes in system strength and others induce changes in inter-system connectivity? Could we create a parameterization of task families that would allow us to manipulate these two phenotypes smoothly and continuously?

What induces network reconfiguration? Who is able to respond to training with greater network reconfiguration and why? What constraints determine what sorts of reconfiguration are easier or harder than others? How much energy does it take to induce a network reconfiguration?


Urs Braun


Mason Porter


Marcelo Mattar


Peter Mucha


Rick F. Betzel


Scott Grafton


Raphael Gerraty


Daphna Shohamy

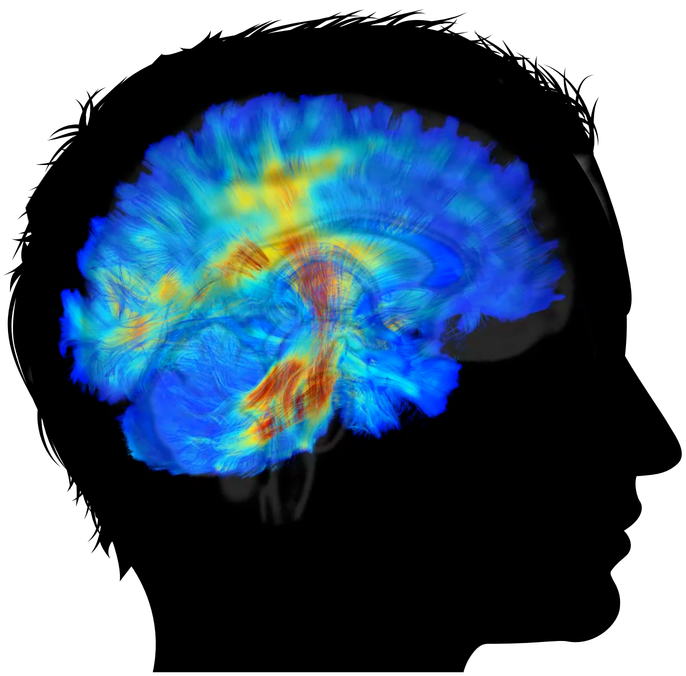
Karuza et al. 2016 *Trends in Cognitive Science*

"Now if there was a becoming of every changeable thing, it follows that before the motion in question another change must have taken place in which that which is capable of being changed or of causing change had its becoming."

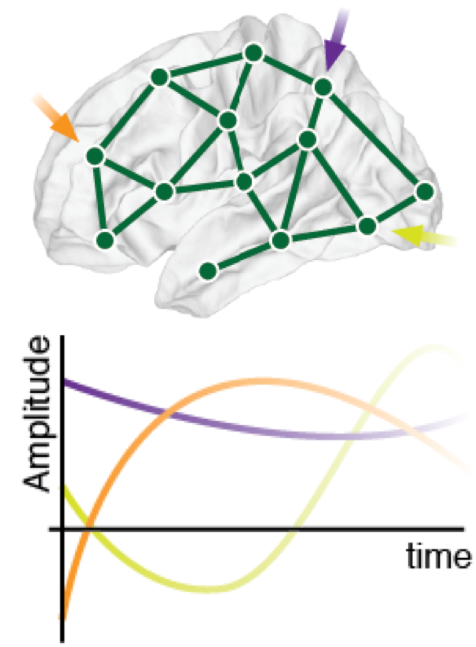
Aristotle, *Physics* VIII.I, 251a9

Constraining Nature of Network Architecture

Can build a theoretical model from data that predicts the changing, the becoming, and the causing of change? Or … how the brain's activity can be altered by a perturbative signal?



What we have: A network of structural links empirically measured by neuroimaging.



What we seek: A theory for how a change in activity in one region affects activity in other regions.

Formalizing the Problem of Network Control

- Neural processes can be approximated by linearized generalizations of nonlinear models of cortical circuit activity (Galan 2008; Honey et al. 2009).
- We consider a noise-free linear discrete-time and time-invariant network model:

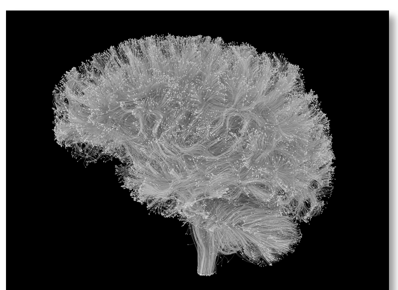
Number of regions being controlled

$$x(t+1) = Ax(t) + B_{\mathcal{K}} u_{\mathcal{K}}(t)$$

State of brain regions over time

Weighted adjacency matrix

Control energy



Gu et al. (2015) *Nature Communications*; Tang et al. (2017) *Nature Communications*

Danielle S. Bassett

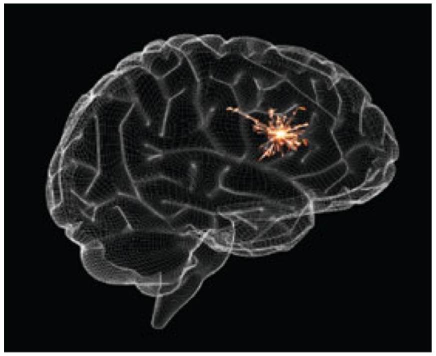
Is the brain theoretically controllable?

How controllable the network is can be estimated using the smallest eigenvalues of the T-steps controllability Gramian:

$$W_{\mathcal{K},T} = \sum_{\tau=0}^{T-1} A^{\tau} B_{\mathcal{K}} B_{\mathcal{K}}^{\mathsf{T}} (A^{\mathsf{T}})^{\tau}$$

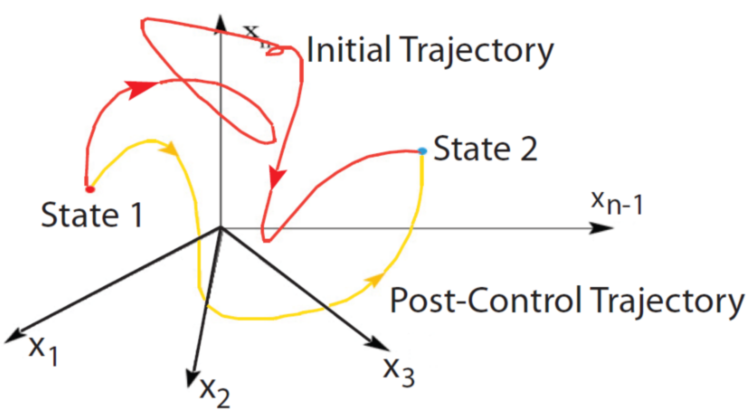
For brain networks, this value was small: 2.5 x 10^(-23)

- Practically extremely hard to control



Types of driver nodes

- Which regions of the brain are most efficient or most difficult to control?

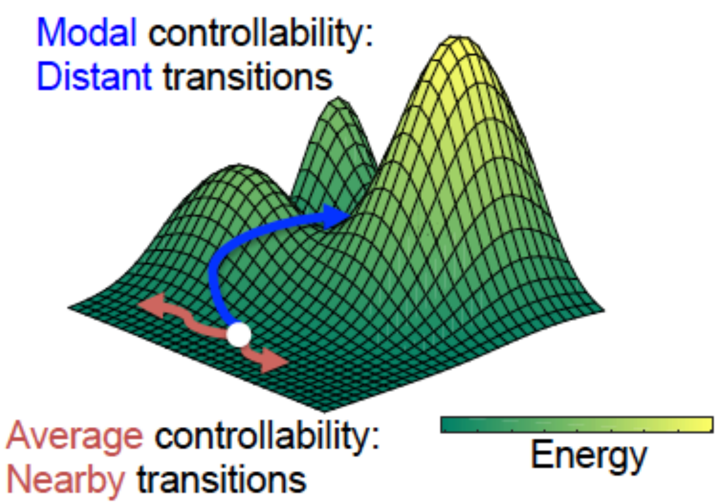


A couple control strategies:

1. **Average Controllability:** Steer to many easily reachable states

2. **Modal Controllability:** Steer to few difficult to reach states

Average and modal control



$$x(t+1) = Ax(t) + B_{\mathcal{K}} u_{\mathcal{K}}(t)$$

$$W_{\mathcal{K},T} = \sum_{\tau=0}^{T-1} A^{\tau} B_{\mathcal{K}} B_{\mathcal{K}}^{\mathsf{T}} (A^{\mathsf{T}})^{\tau}$$

Average: Trace(W_K^{-1}))

Modal: Let v_j be the j^{th} eigenvector of A with eigenvalue λ_j. Then if v_{ij} is small, then the j^{th} mode is poorly controllable from node i. Define $\phi_i = \sum_{j=1}^{N} (1 - \lambda_j^2(A)) v_{ij}^2$ as a scaled measure of controllability of all N modes from region i.)

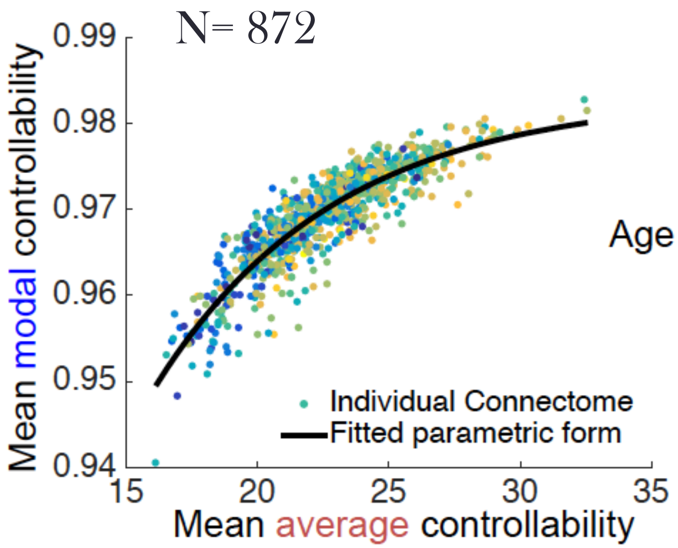
Network control as a model for cognitive control

- Different brain regions have more or less average/modal controllability, indicating differential capacity to alter whole-brain dynamics
  (Gu et al. 2015 *Nature Communications*)

- The capacity of brain regions to exert control grows as children develop
  (Tang et al. 2017 *Nature Communications*)

- Network controllability is correlated with impulsivity, a measure of executive function
  (Cornblath et al. 2018 NeuroImage)

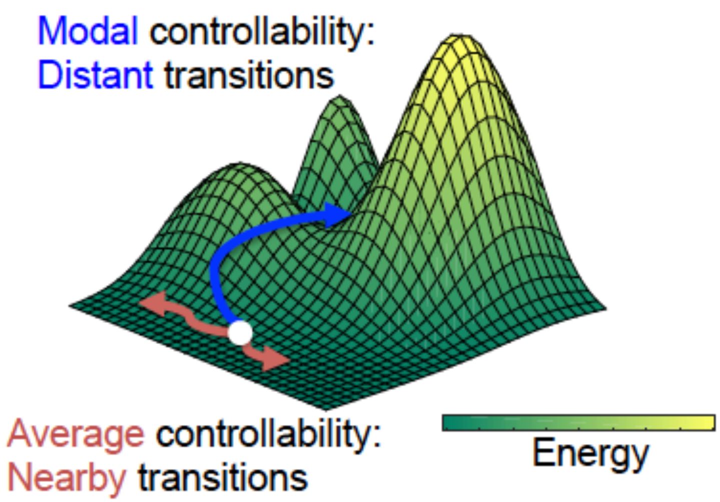
- The energy required for control decreases with age, in concert with increasing executive function
  (Ciu et al. 2018 In Revision; preprint available on BioRxiv)

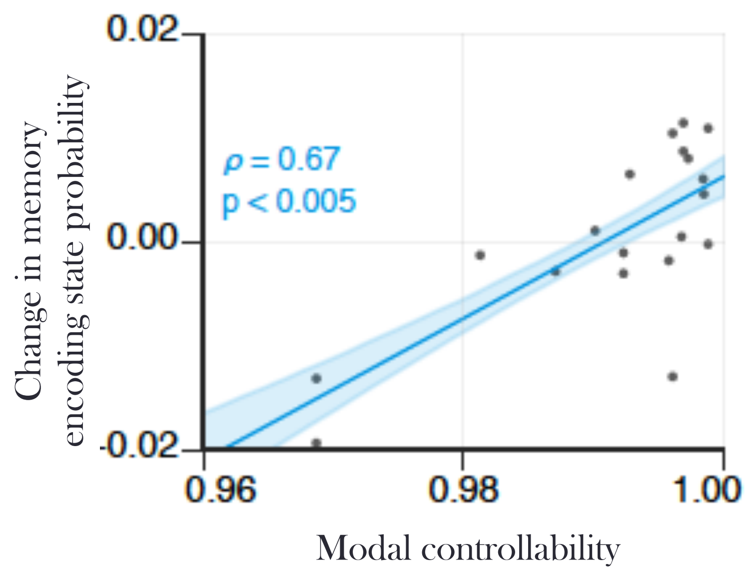


Together, these results suggest that our theory is a useful marker of how the brain enacts control to change network function.

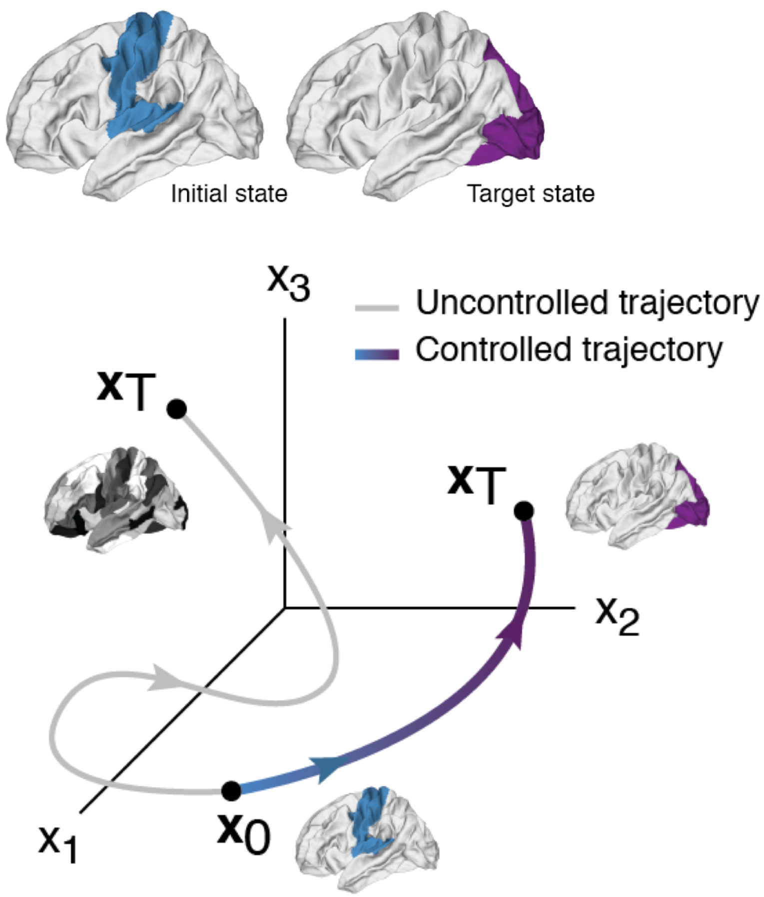
Extending to exogenous control (stimulation)

Preliminary work suggests that stimulation to modal controllers pushes the brain into better memory encoding states.





Precise control of specific state transitions



What we want

- Finite time, Finite energy,
- Multi-point control
- Initial state, Target state

Define model of network dynamics.

$$x(t+1) = Ax(t) + B_{\mathcal{K}} u_{\mathcal{K}}(t)$$

Define a cost function penalizes energy and distance of x(t) from the target state.

$$\min_{\mathbf{u}} \int_0^T (\mathbf{x}_T - \mathbf{x})^T (\mathbf{x}_T - \mathbf{x}) + \rho \mathbf{u}_{\mathcal{K}}^T \mathbf{u}_{\mathcal{K}}$$

Betzel et al. (2016) *Scientific Reports*; Gu et al. (2017) *Neuroimage*
Stiso et al. (2019) In Press at Cell Reports; coming out Sept 3

Open questions

What is it about certain network topologies that makes them easier or harder to control? (Kim et al. 2018 *Nature Physics*) Does the answer to this question help us to understand time scales of control, such as transient versus persistent control, which may be altered in certain patient groups? (Tang et al. 2019, Phys Rev E, In Revision)

How does control of brain state transitions relate to network reconfiguration exactly? How might brain states relate to neural representations, housing information about the world or our model of the world? What rules constrain the evolution of neural representations during learning? (Tang et al. (2019) *Nature Neuroscience*)

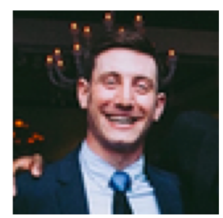

Fabio Pasqualetti


Jason Z. Kim


Ankit Khambhati

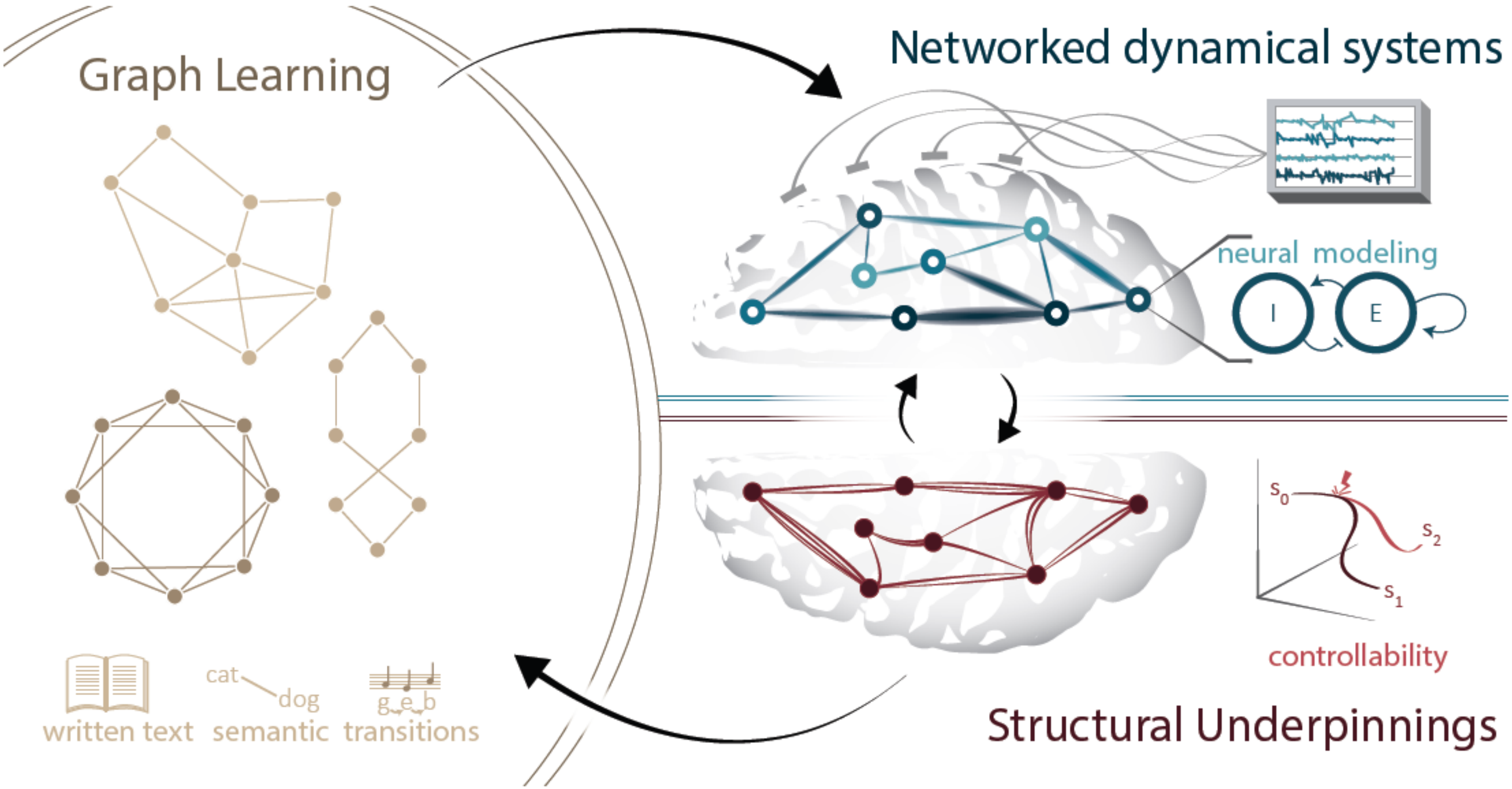

Shi Gu


Rick F. Betzel

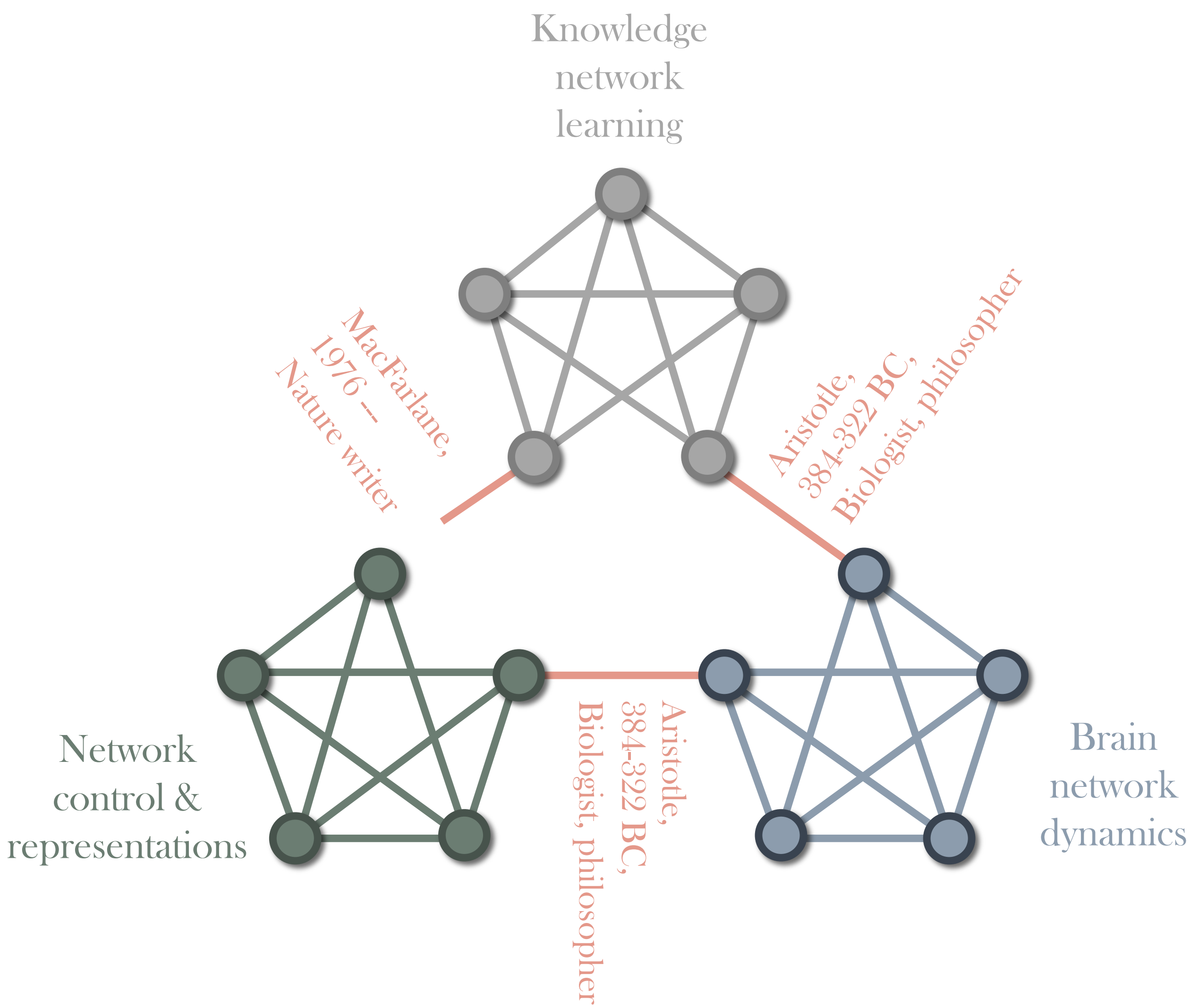

Eli Cornblath


Jeni Stiso

Summary & Future Directions



Networks thinking themselves



Acknowledgments

TEAM: Abigail Poteshman, Andrew Murphy, Ann Sizemore-Blevins, Ari Kahn, Arun Mahadevan, Cedric Xia, Christopher Lynn, Dale Zhou, David Lydon-Staley, Eli Cornblath, Erin Teich, Graham Baum, Harang Ju, Jason Kim, Jeni Stiso, Karol Szymula, Keith Wiley, Lia Papadopolous, Lindsay Smith, Lorenzo Caciagli, Mark Choi, Maxwell Bertolero, Pranav Reddy, Urs Braun, Ursula Tooley, Xiaosong He, Zhixin Lu.

Past trainees now faculty:


Sarah Muldoon


Shi Gu

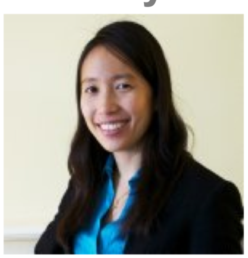

John Medaglia


Ralf Schmaezle


Chad Giusti

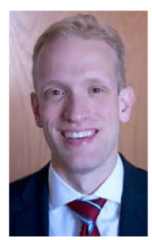
Brand new faculty members:


Lizz Karuza


Evelyn Tang


Rick F. Betzel

Faculty collaborators:


Ted Satterthwaite


Raquel Gur


Fabio Pasqualetti

Networks thinking themselves