Testing a Common Model for Human and Machine Intelligence

(FA9550-19-1-0299)

PI: Andrea Stocco (Univ. of Washington)

Co-PIs: John Laird (Univ. of Michigan)

Christian Lebiere (CMU)

Paul Rosenbloom (USC)

Senior Personnel: Catherine Sibert

AFOSR Program Review:
Computational Cognition and Machine Intelligence Program
(October 6-8, 2020, virtual program review)

Testing a Common Model for Human and Human-Like Intelligence

Andrea Stocco, University of Washington

Objective:

- Develop principles to translate cognitive architectures into dynamical brain network models
- Test the reliability of Common Model of Cognition (CMC, a consensus architecture) as a blueprint of brain architecture

Approach:

- Compare alternative network models by fitting them to large-scale neuroimaging data
- HCP: Human Connectome Project (fMRI + MEG)

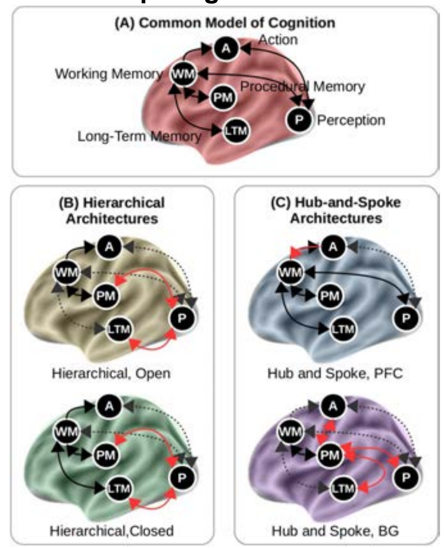
DoD Benefits:

- Link between cognitive architectures and computational neuroscience
- Convergent evolution of human and human-like intelligence
- General-purpose neurally-inspired architecture for AI

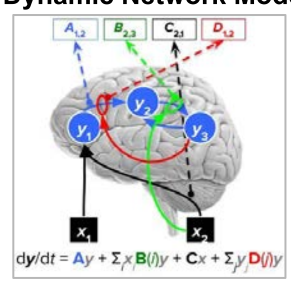
Progress:

- Tested fMRI data from *N*=200 participants from HCP dataset performing 7 tasks (Social, Emotion, WM, DM, Language, Math, Reasoning)
- Publications: Two journal papers + one conference paper submitted
- New pipeline in development: Bilateral ROIs, parallel processing, resting state data

1. Competing Architectures

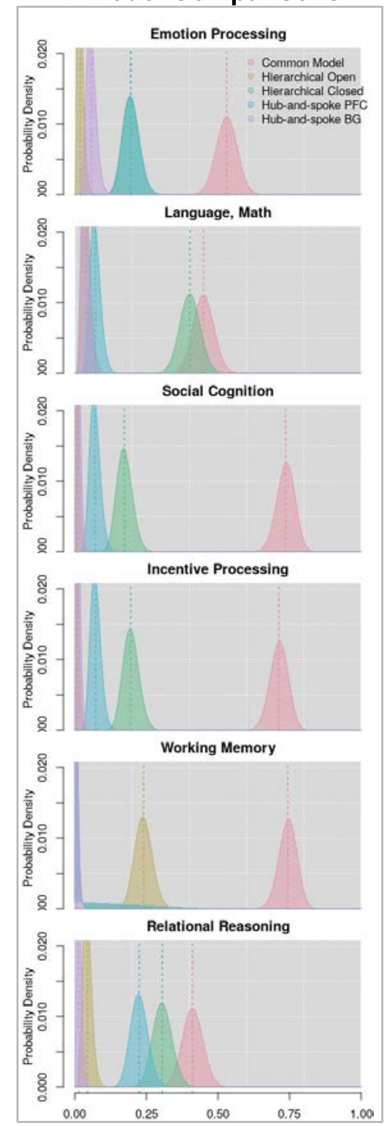


2. Dynamic Network Models



3. fMRI Data (*N* = 200, 7 Tasks)

4. Model Comparisons



List of Project Goals

1. Develop a **methodology** to translate cognitive architectures into testable network brain architectures
   - Language for abstract architecture definition
   - Pipeline to map components to brain and fit models
2. Test and compare **alternative architectures**
   - Task-based fMRI data
   - Resting-state data
3. **Validation** of the results
   - Parameter analysis
   - Alternative methods
   - MEG/EEG data

Progress Towards Goals (or New Goals)

1. Develop a **methodology** to translate cognitive architectures into testable network brain architectures
   - Language for abstract architecture definition
   - Pipeline to map components to brain and fit models
2. Test and compare **alternative architectures**
   - Task-based fMRI data
   - Resting-state data
3. **Validation** of the results
   - Parameter analysis
   - Alternative methods
   - MEG/EEG data
   - Ground Truth with Clinical Population

A Standard Model of the Mind: Toward a Common Computational Framework Across Artificial Intelligence, Cognitive Science, Neuroscience, and Robotics

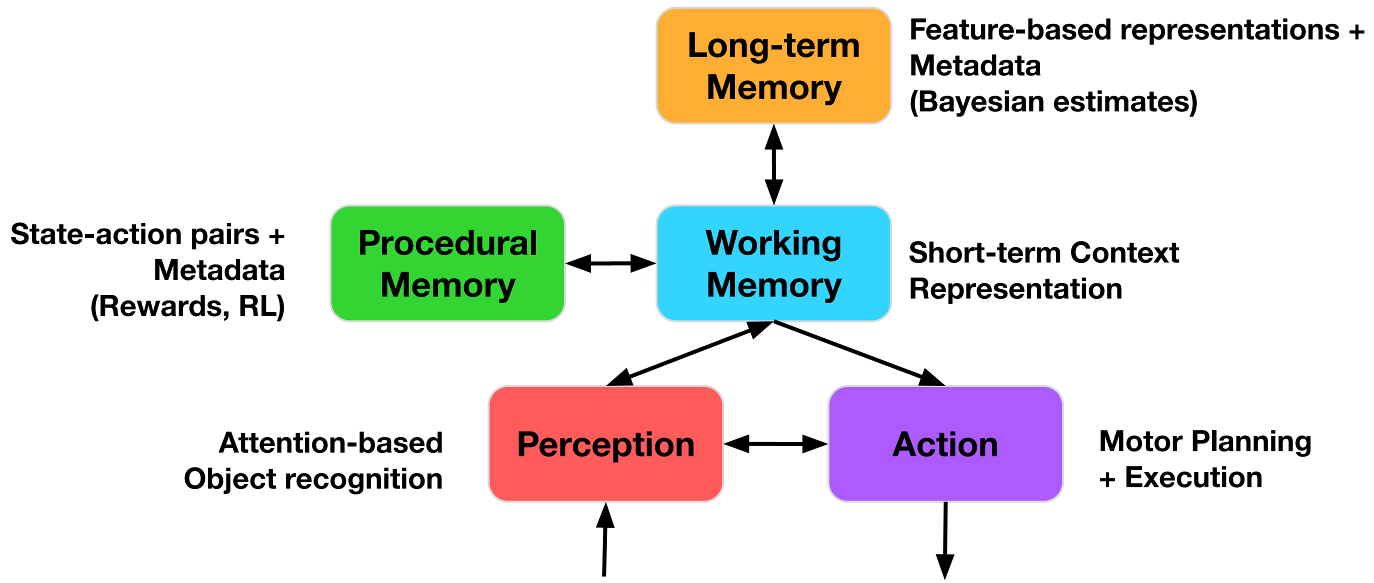
John E. Laird, Christian Lebiere, Paul S. Rosenbloom

■ *A standard model captures a community consensus over a coherent region of science, serving as a cumulative reference point for the field that can provide guidance for both research and applications, while also focusing efforts to extend or revise it. Here we propose developing such a model for humanlike minds, computational entities whose structures and processes are substantially similar to those found in human cognition. Our hypothesis is that cognitive architectures provide the appropriate computational abstraction for defining a standard model, although the*

A mind is a functional entity that can think, and thus support intelligent behavior. Humans possess minds, as do many other animals. In natural systems such as these, minds are implemented through brains, one particular class of physical device. However, a key foundational hypothesis in artificial intelligence is that minds are computational entities of a special sort — that is, cognitive systems — that can be implemented through a diversity of physical devices (a concept lately reframed as substrate independence [Bostrom 2003]), whether natural brains, traditional general-purpose computers, or other sufficiently functional forms of hardware or wetware.

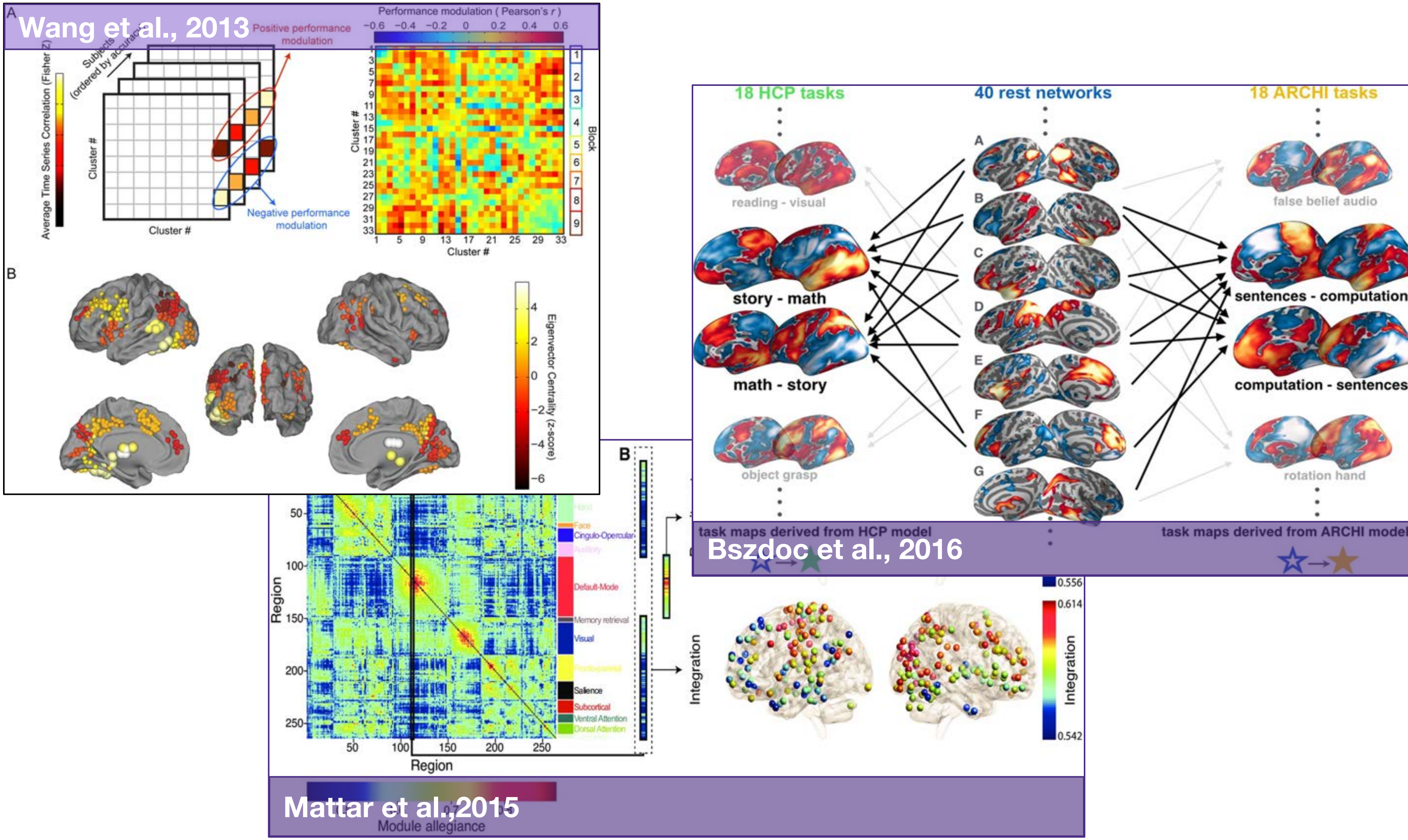
Laird, Lebiere, & Rosenbloom, 2017, AI Mag

Common Model of Cognition



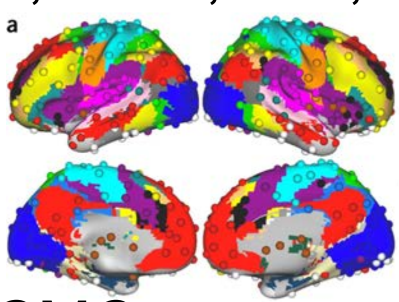
Could the CMC be a candidate **brain** architecture?

Could the CMC be a candidate **brain** architecture?

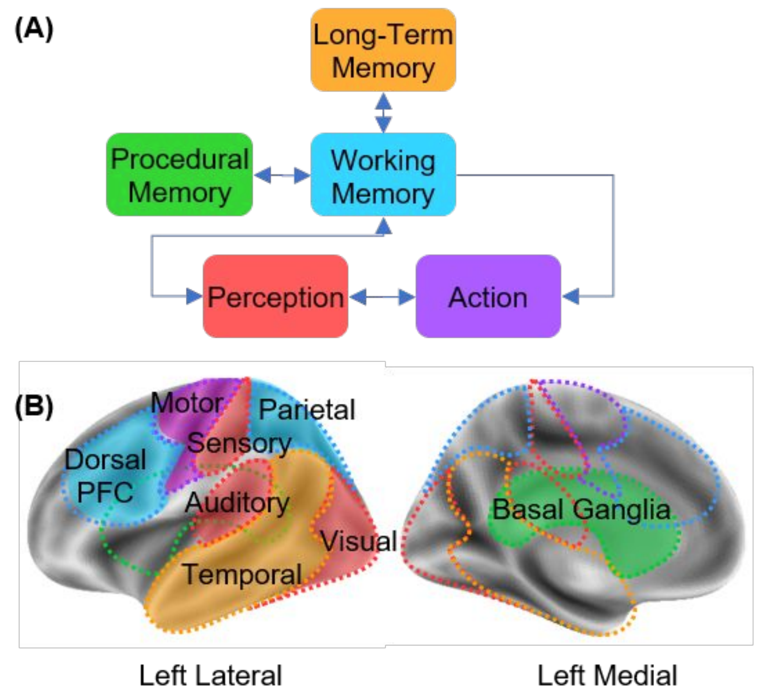


Mapping CMC Components

- Human brain is divided into small number (~ 10) of **functional networks** (Power, 2010; Yeo, 2011)



- Each CMC components can be identified with **at least one.**



Data: Human Connectome Project

- Contains high-quality neuroimaging data:
  - 1,200 Adult Participants
  - 7 Different Tasks
  - 4 Resting State Sessions
  - fMRI + MEG data (subset)

Data: Human Connectome Project

- Contains high-quality neuroimaging data:
  - **200 out of** 1,200 **Adult Participants**
  - **6** out of 7 **Different Tasks**
  - 4 Resting State Sessions
  - **fMRI** + MEG data (subset)

Siemens Skyra, Multiband					
TR	720 ms	**MB factor**	8x	***N* Slices**	72
TE	33.1 ms	**FOV**	208 x 180 mm	**Slice Gap**	0mm
FA	52°	**In-plane res**	2 x 2 mm	**Slice thick**	2mm

What Tasks?

Task	Reference	Description
Motor	Buckner et al. (2011)	**Hand, arm, foot, leg, voice** responses
Emotional	Hariri et al. (2002)	**Fearful faces** vs. Neutral Shapes
Incentive Processing	Delgado et al. (2000)	**“Losing” blocks** vs. “Winning” blocks of choices
Language	Binder et al. (2011)	**Language blocks** vs. Math blocks
Relational	Smith et al. (2007)	**Relational arrays** vs. Control arrays
Social	Whitley et al. (2007)	**Interacting shapes** vs. Randomly moving
Working Memory	Dobryshevsky et al. (2006)	**2-back** vs 0-back blocks

Comparison Criteria

We cannot tell whether the CMC is “true”, but we can tell whether it is “better”

- **Superiority**
  - The CMC should provide a better account of brain activity than alternative architectures
- **Generality**
  - The superiority of the CMC should be invariant across tasks

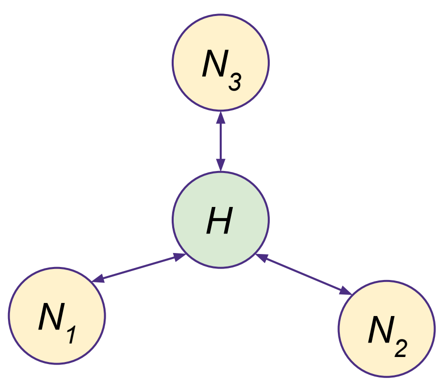
Alternative Architectures

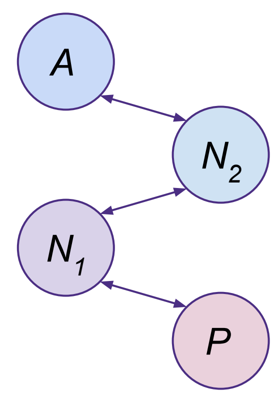
"Hub-and-Spoke" Family

- Single central region, *H*
- Multiple nodes N_1, N_2, N_3
- Bidirectional connections (spokes) btw hub and nodes

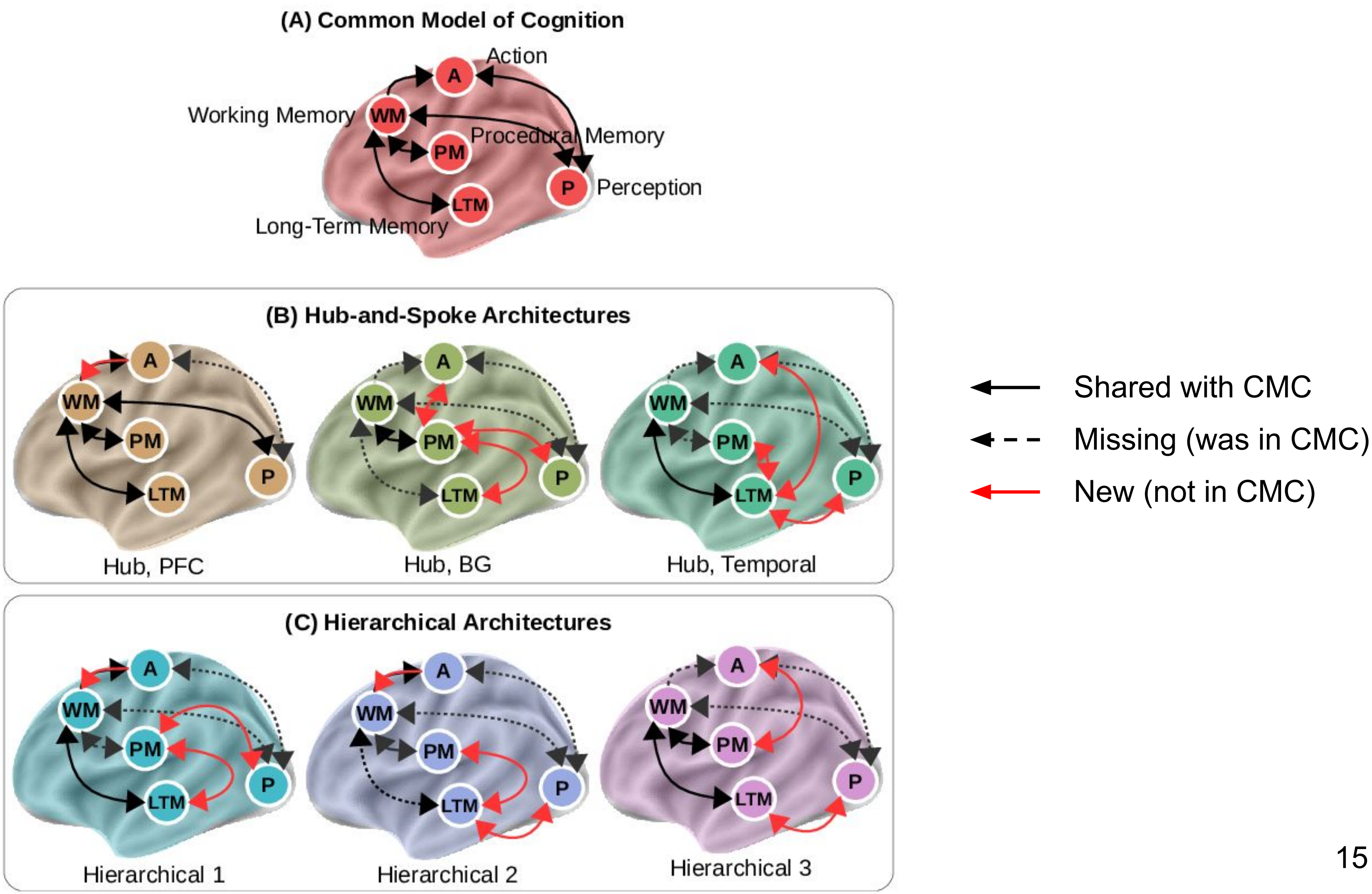
"Hierarchical" Family

- Starts with Perception (*P*)
- Ends with Action (*A*)
- Bidirectional connections between consecutive nodes



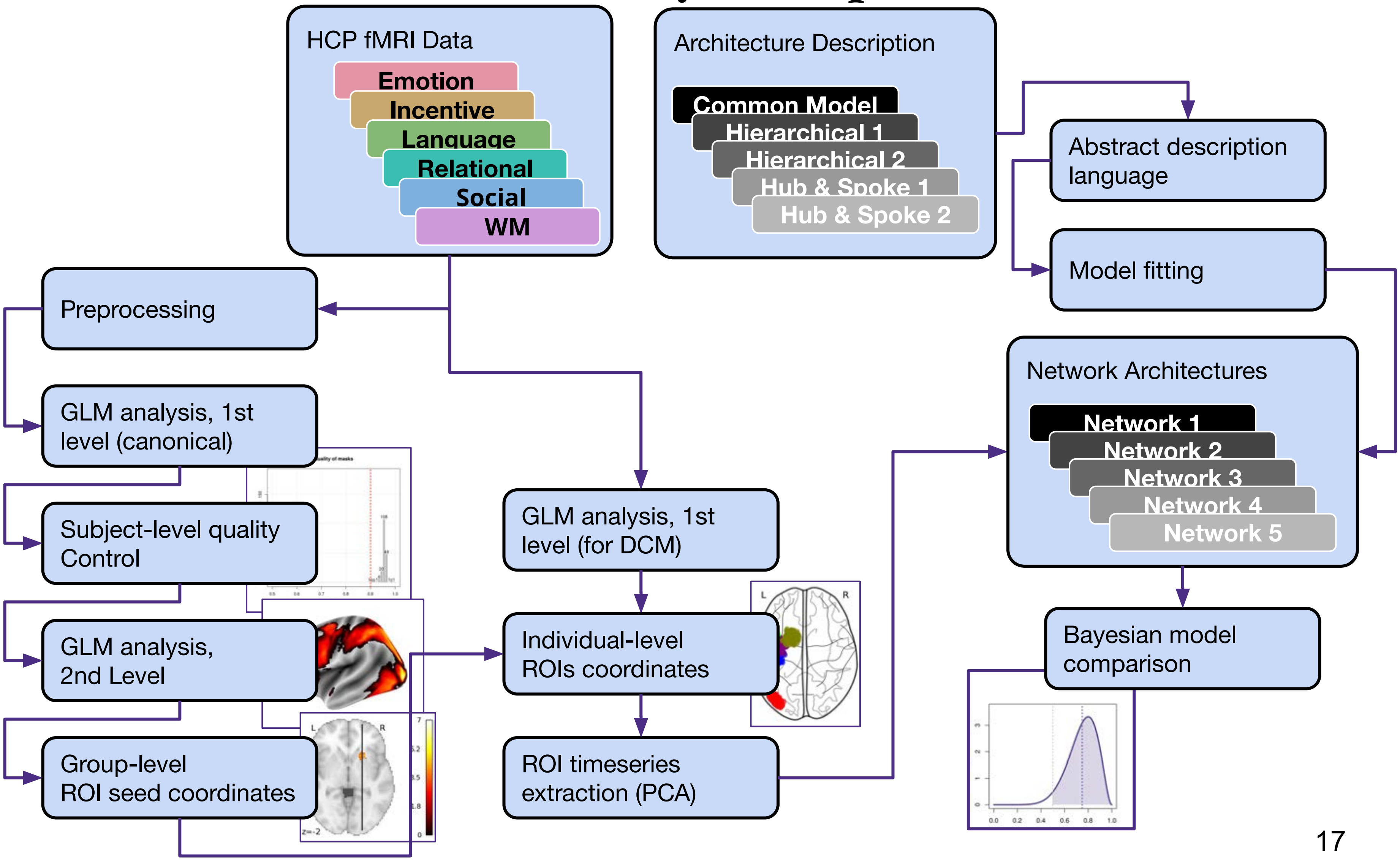


Seven Alternative Architectures

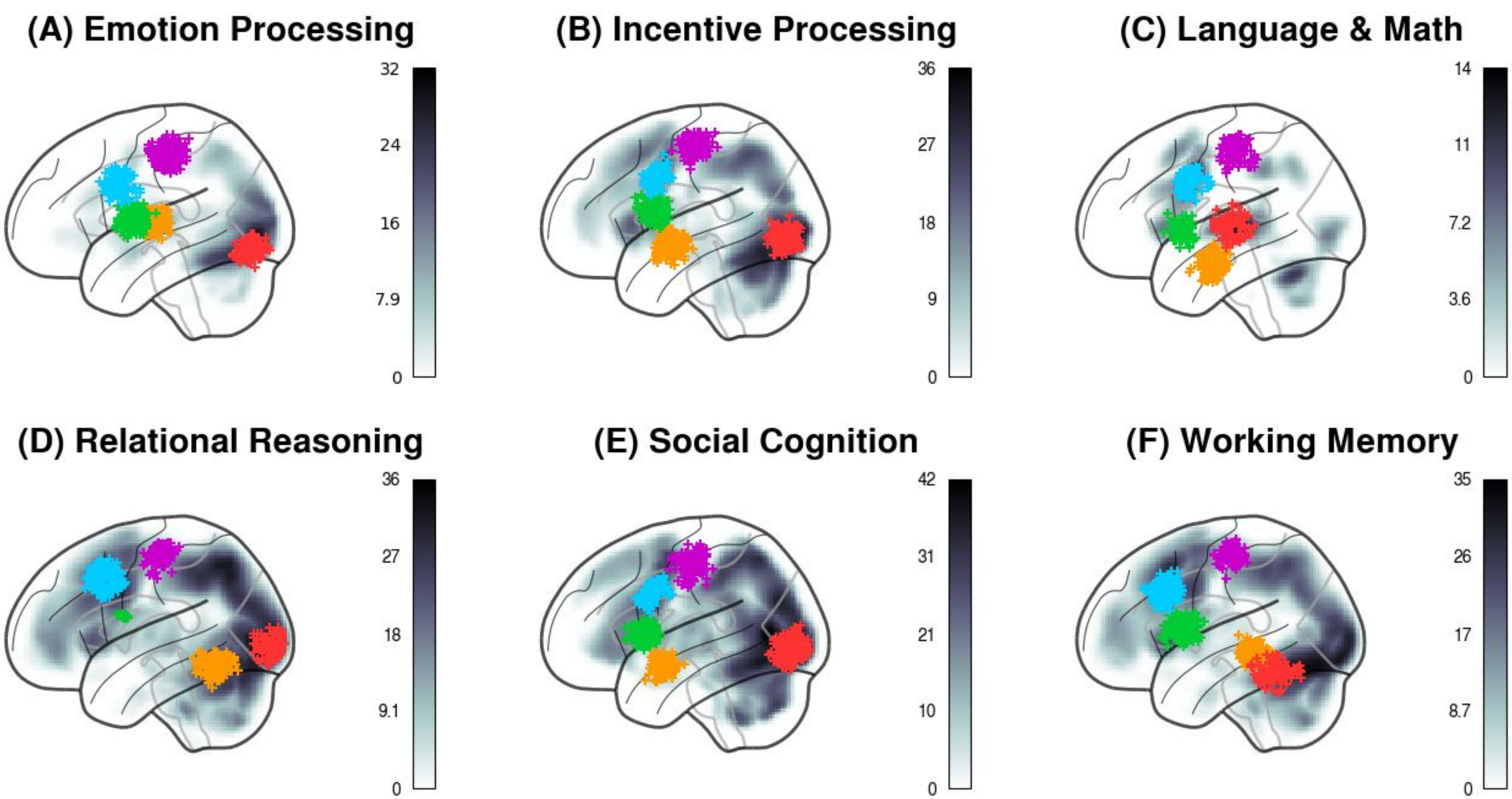


“Fantastic Architectures and Where to Find Them”

Data Analysis Pipeline

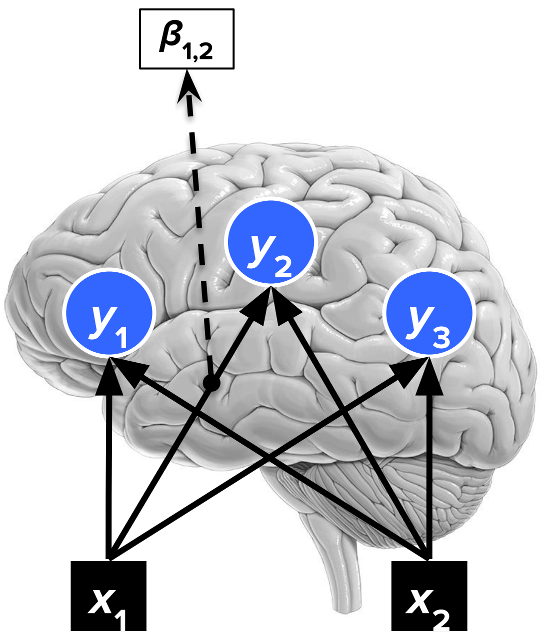


Define the Regions of Interests (ROIs)
For Each Network



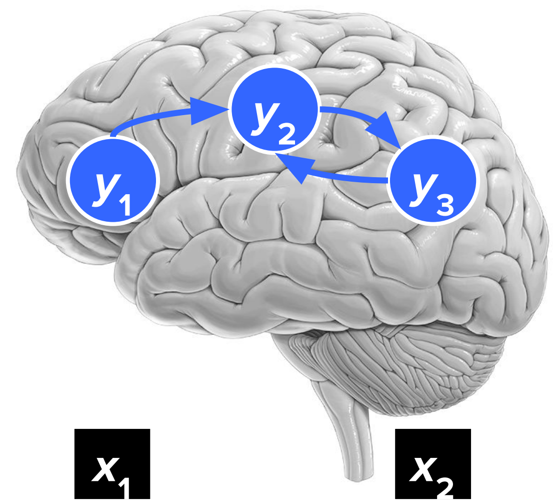
Fitting Network Models

Traditional GLM



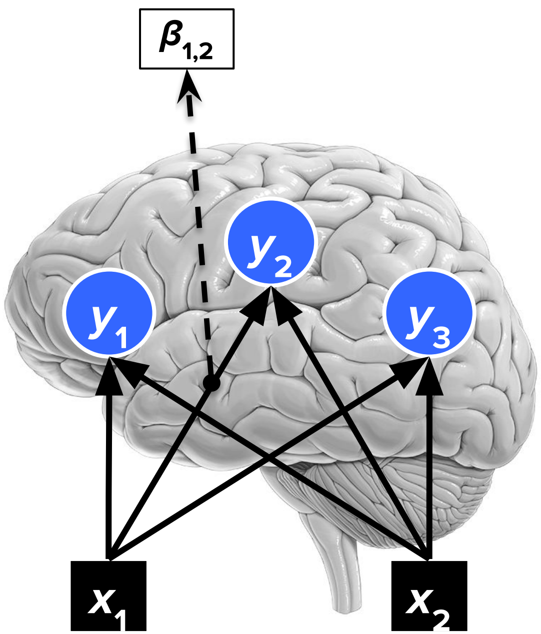
$$\mathbf{y} = \Sigma_i \, \boldsymbol{\beta}_i{}^{*} \mathbf{x}_i$$

?



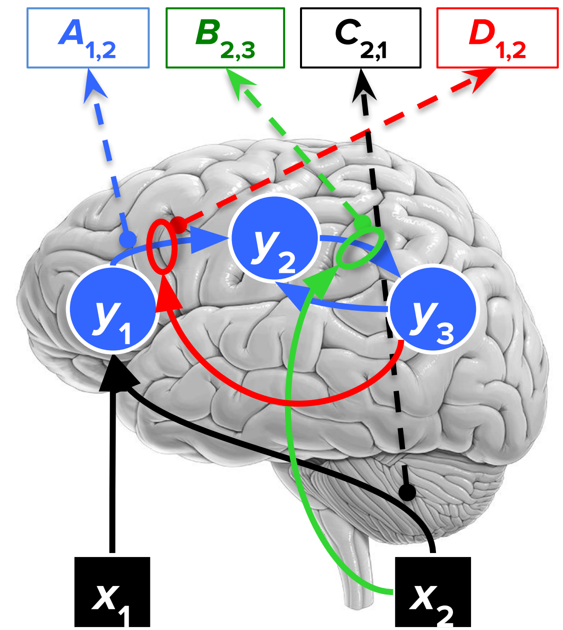
Fitting Network Models

Traditional GLM



$$\mathbf{y} = \Sigma_i \boldsymbol{\beta}_i{}^{*}\mathbf{x}_i$$

Dynamic Causal Modeling

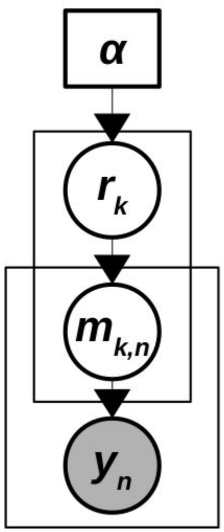


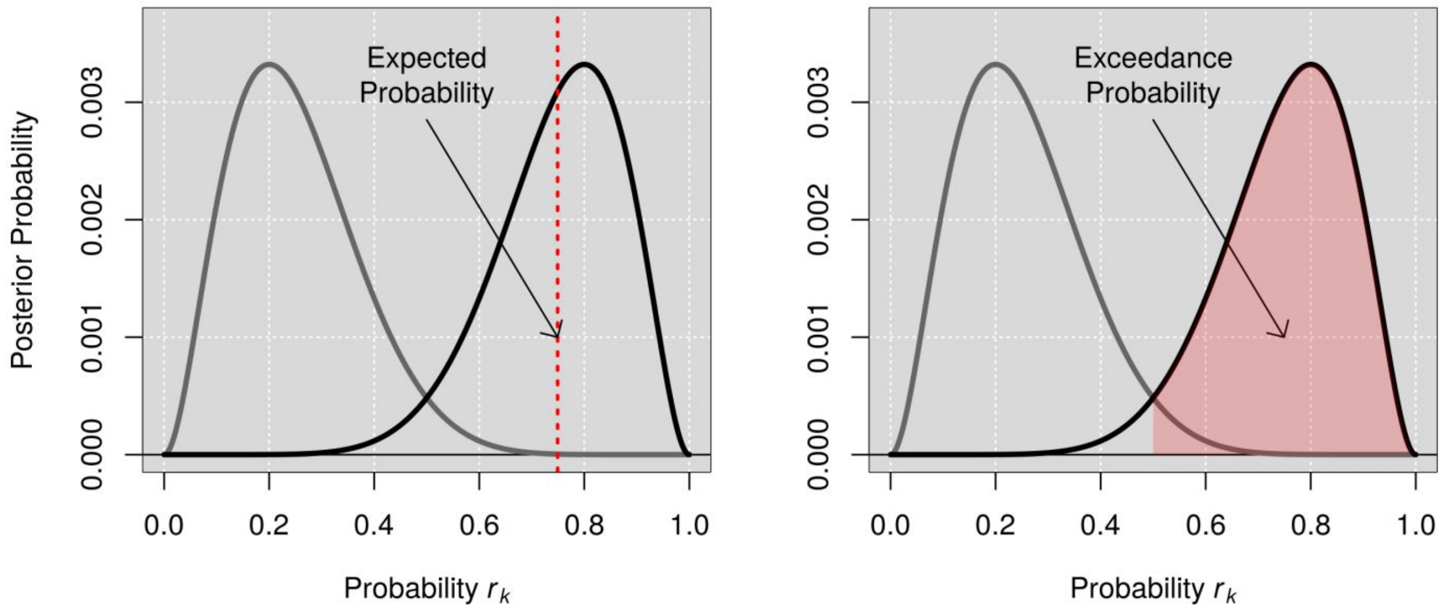
$$\mathrm{d}\boldsymbol{y}/\mathrm{d}t = \mathbf{A}\boldsymbol{y} + \Sigma_i x_i \mathbf{B}(i)\boldsymbol{y} + \mathbf{C}\boldsymbol{x} + \Sigma_j y_j \mathbf{D}(j)\boldsymbol{y}$$

Comparing Models

- Many criteria exists
  - AIC, BIC, Log-likelihood…
- Inter-subject variability a major concern, so we used **Bayesian approach**:
  - Posterior probability that a model is true, given the data
  - Each architectures' PDF is modeled as a Dirichlet distribution ~ Dir(α)

Comparing Models



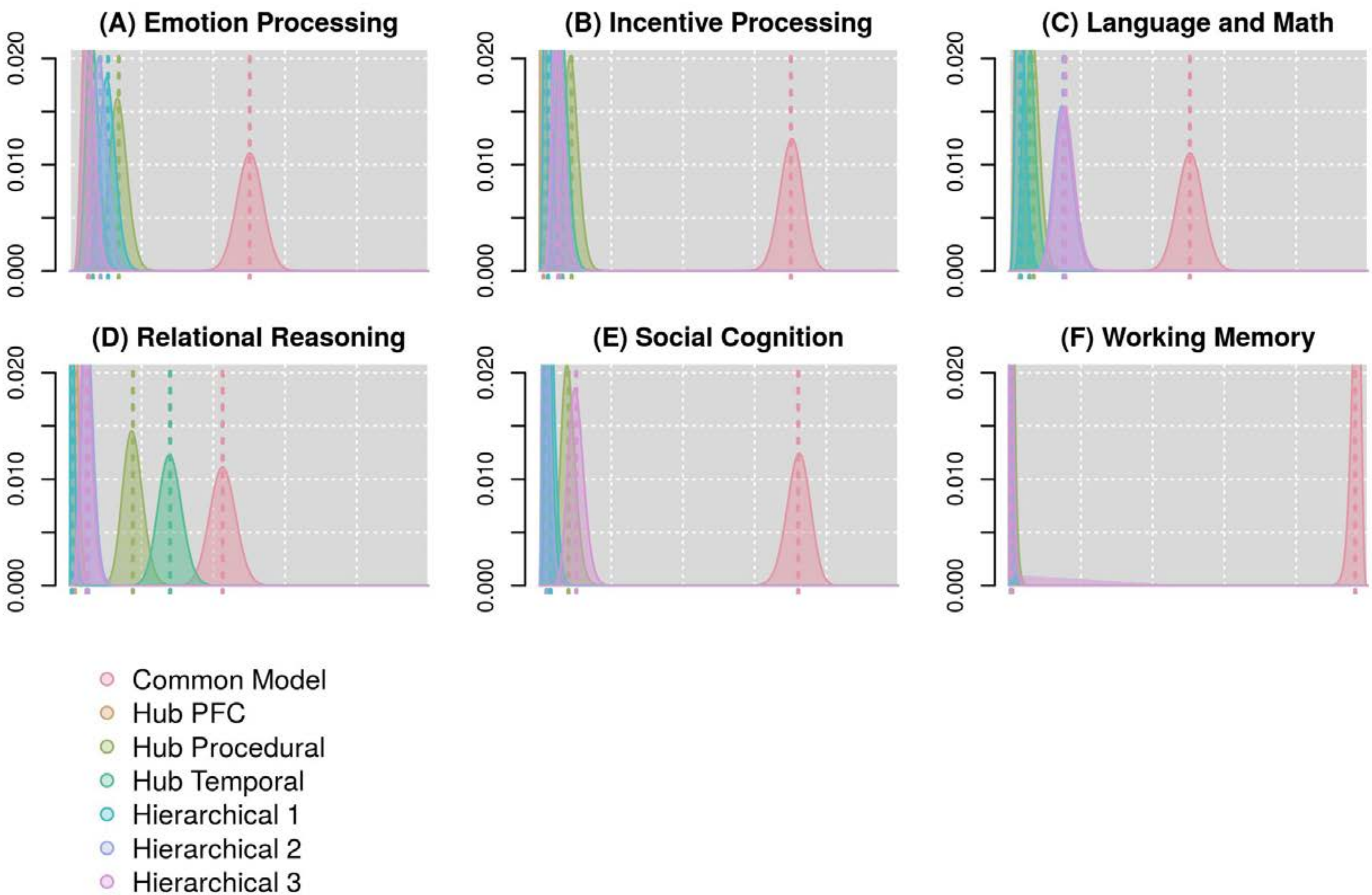


Comparison Criteria

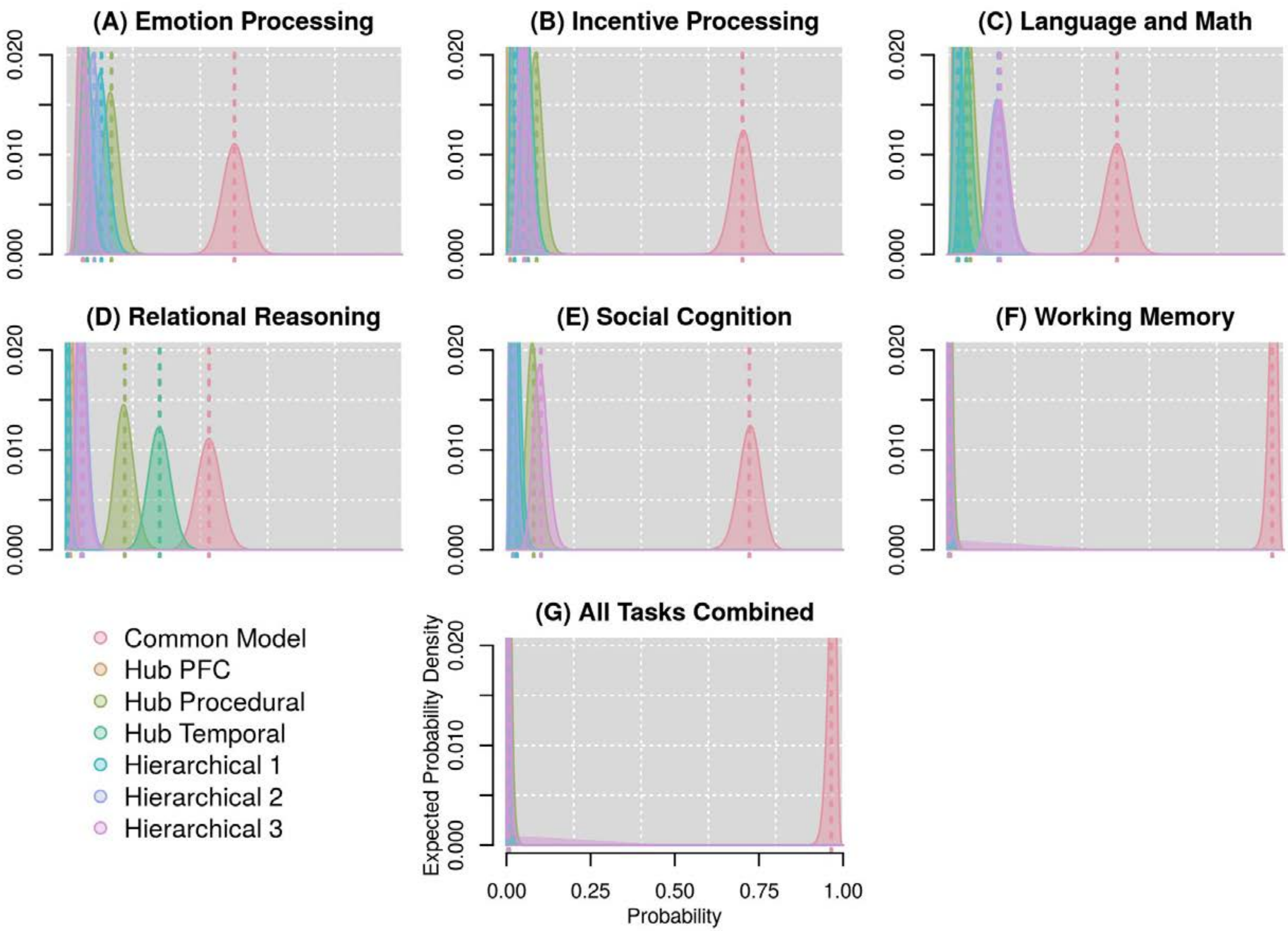
Two criteria

- **Superiority**
  - The CMC should provide a better account of brain activity than alternative architectures
- **Generality**
  - The superiority of the CMC should be invariant across tasks

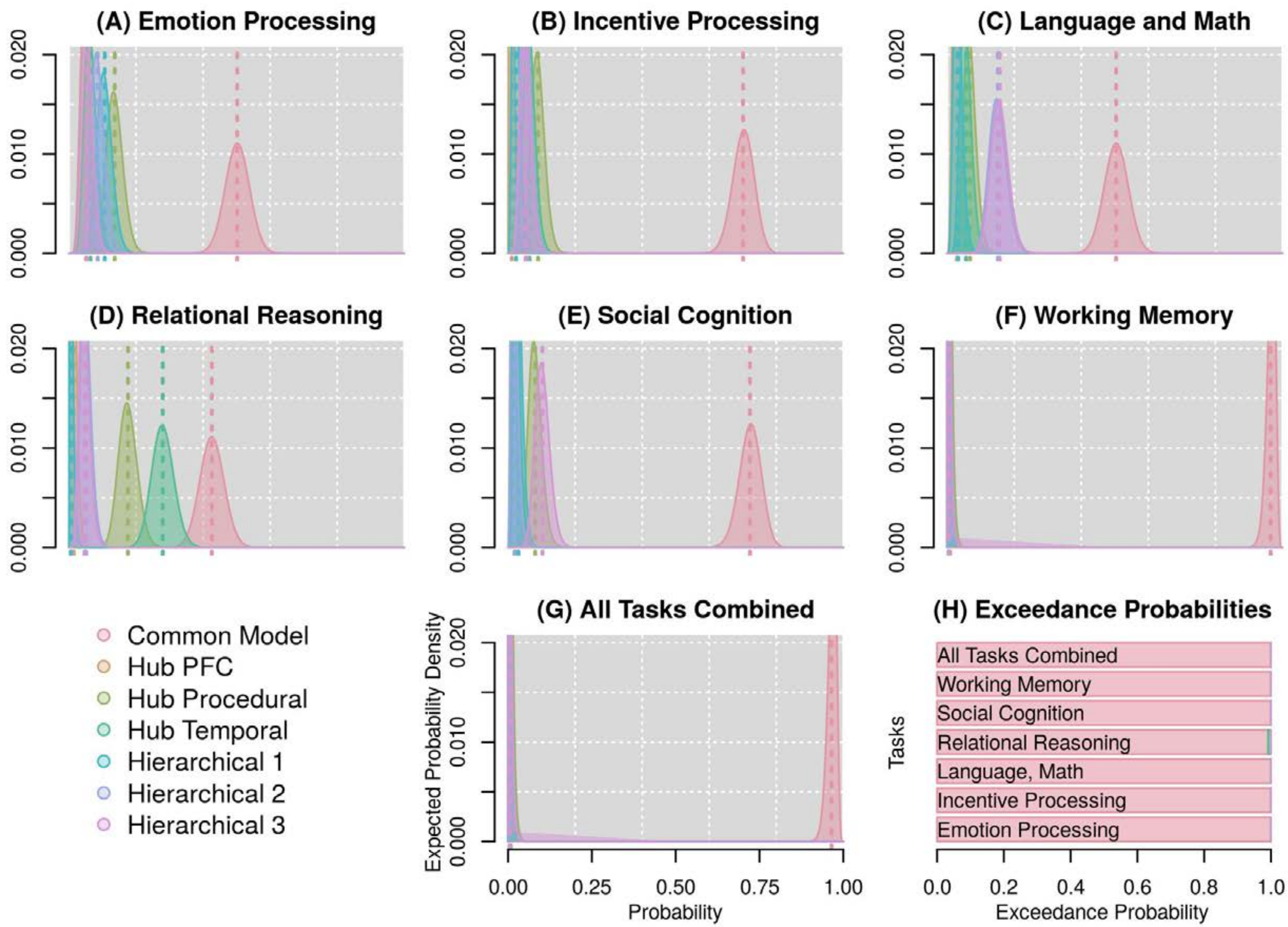
Results: All Tasks



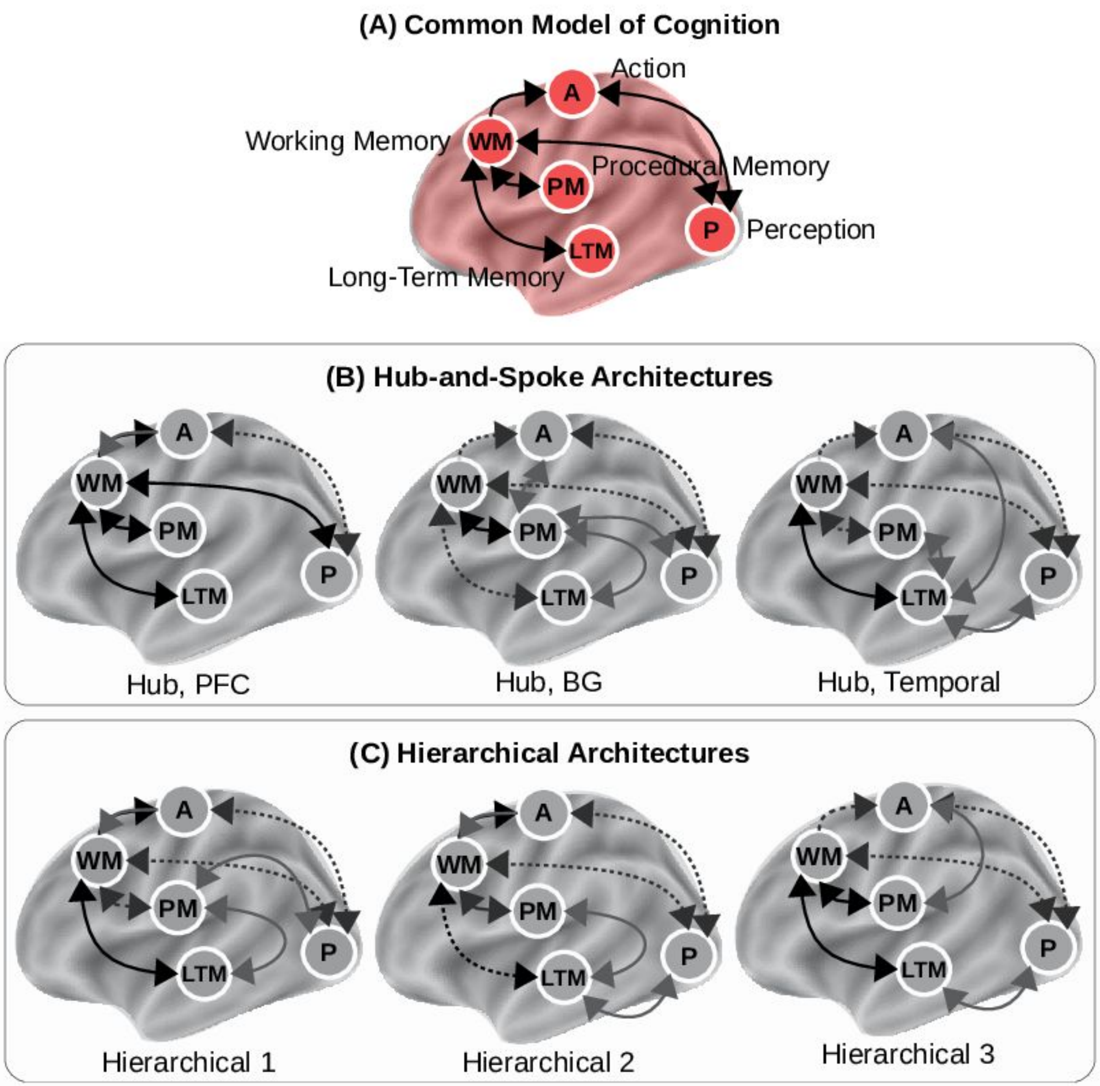
All Tasks, Separate + Combined

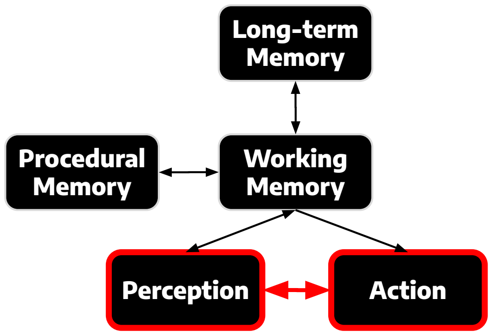


Exceedance Probabilities

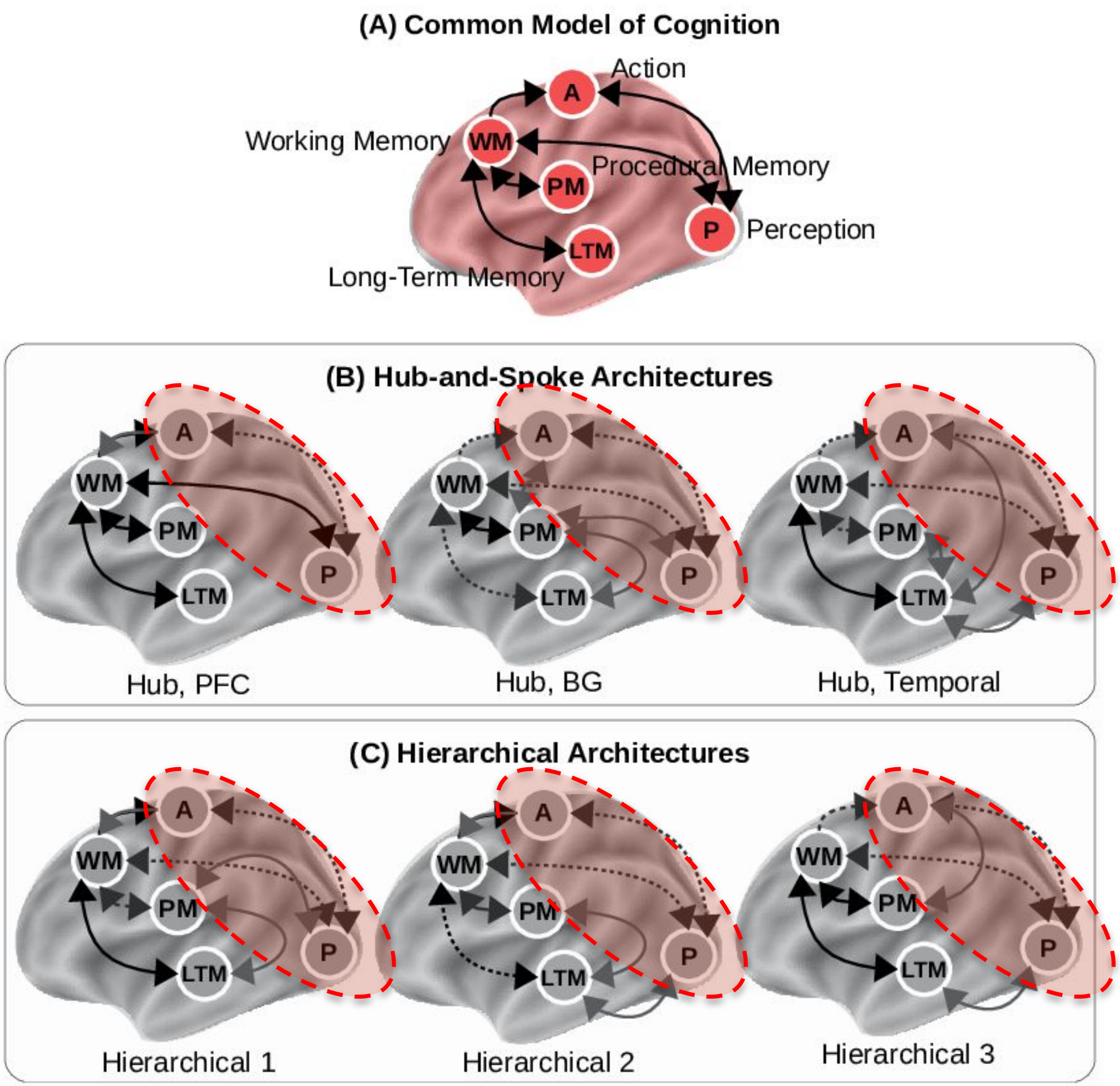


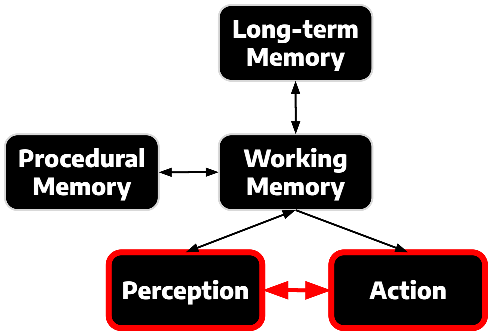
The $P \leftrightarrow A$ Confound



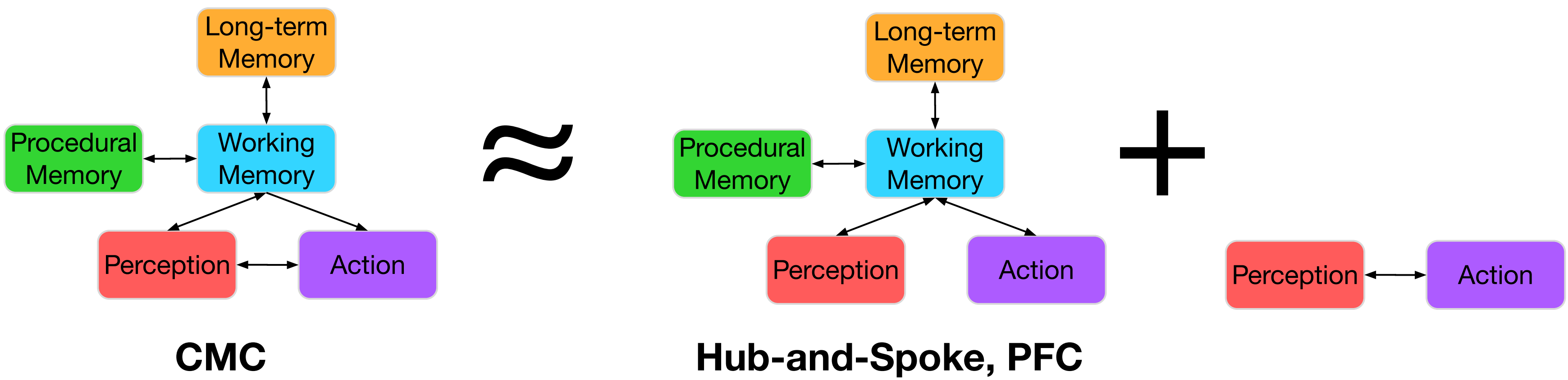


The $P \leftrightarrow A$ Confound



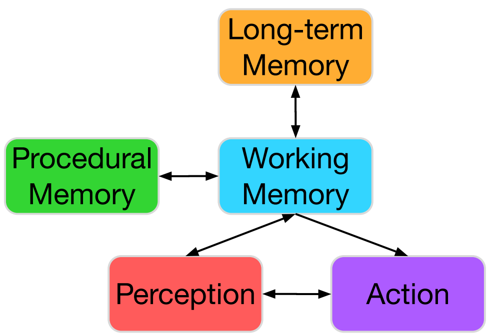


CMC ≈ Hub-and-Spoke PFC + $P \leftrightarrow A$



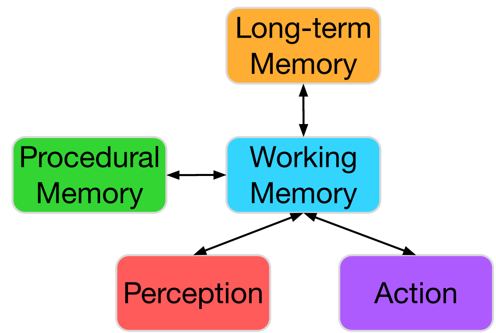
CMC ≈ Hub-and-Spoke PFC + $P \leftrightarrow A$

CMC

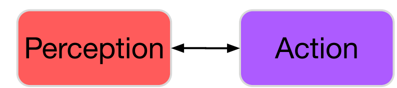


≈

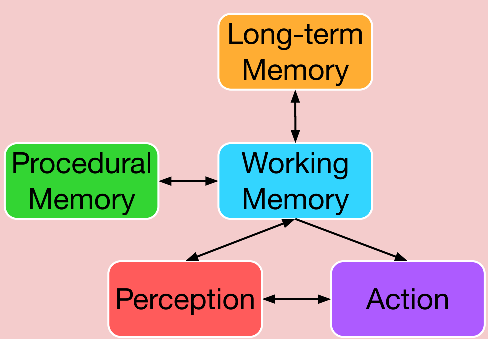
Hub-and-Spoke, PFC

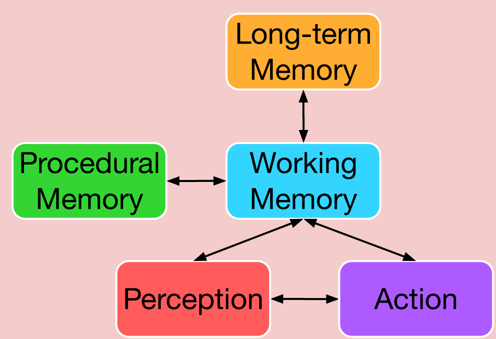


+



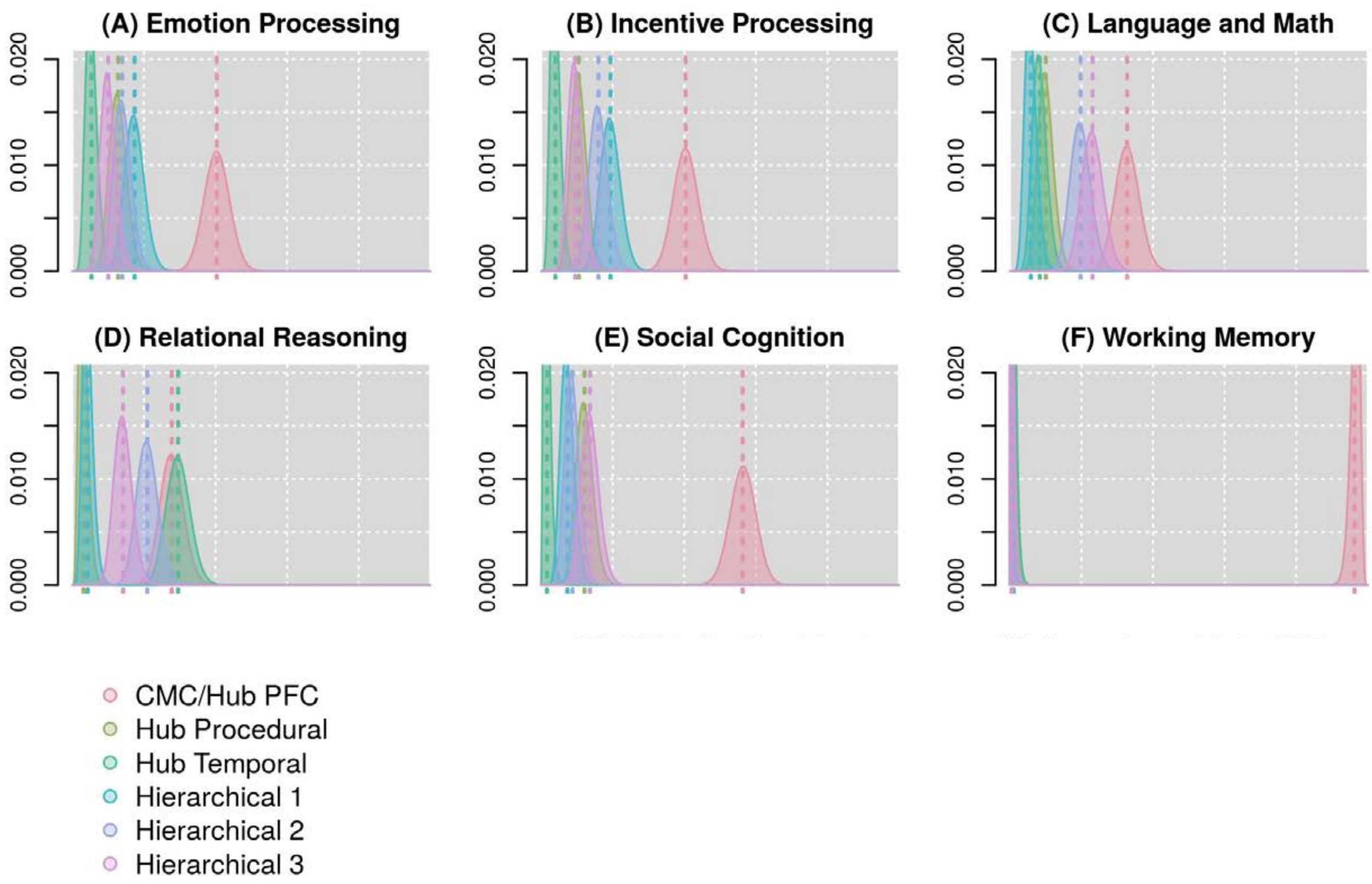
CMC/Hub PFC Family



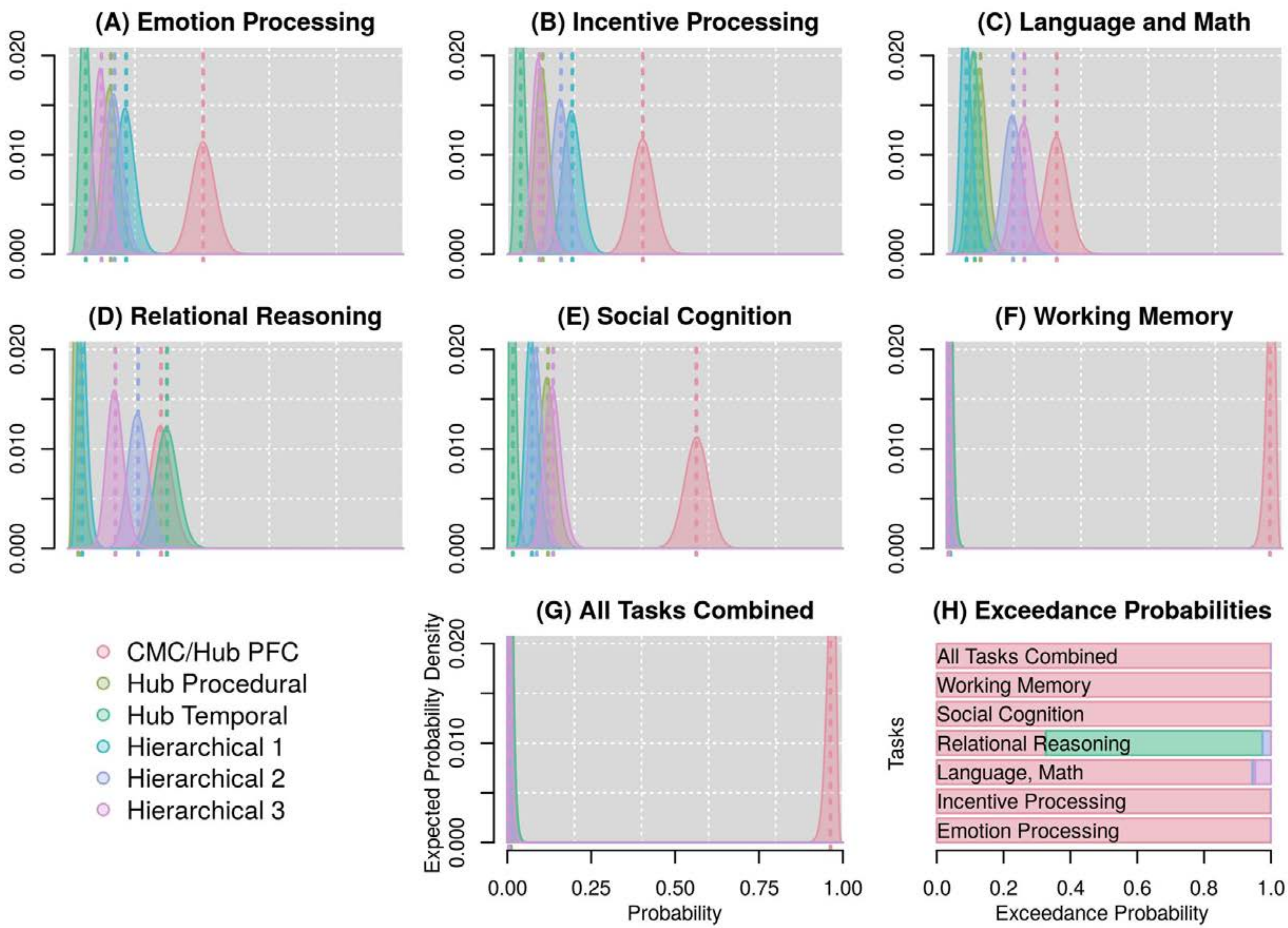


CMC **Hub-and-Spoke, PFC**

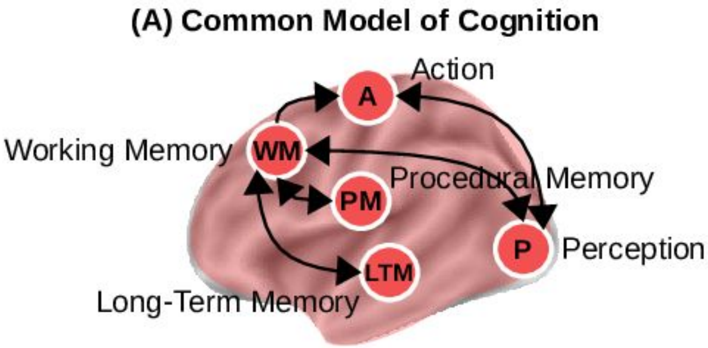
Results, All Architectures + $P \leftrightarrow A$



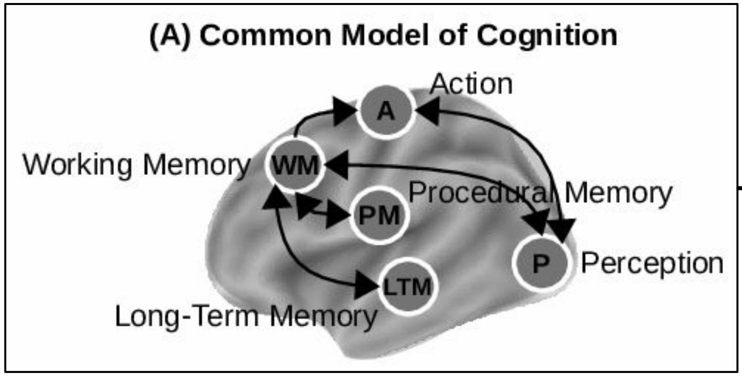
Results, All Architectures + $P \leftrightarrow A$

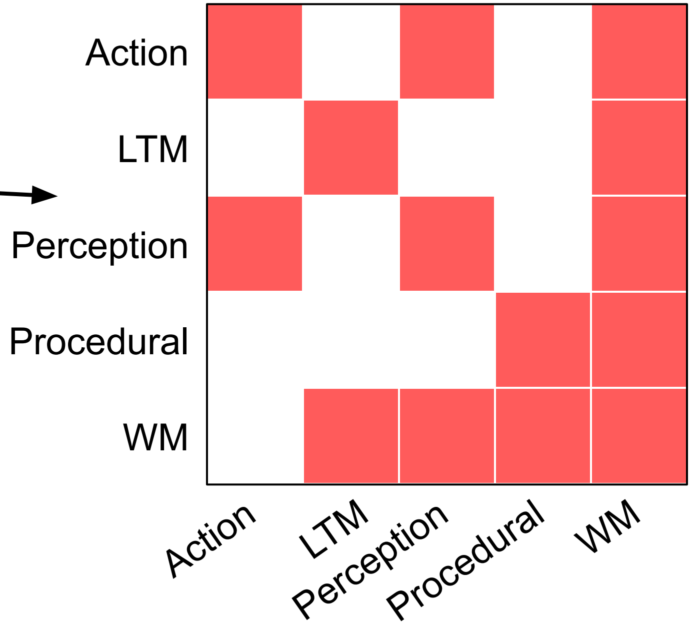


Are All Connections Necessary?

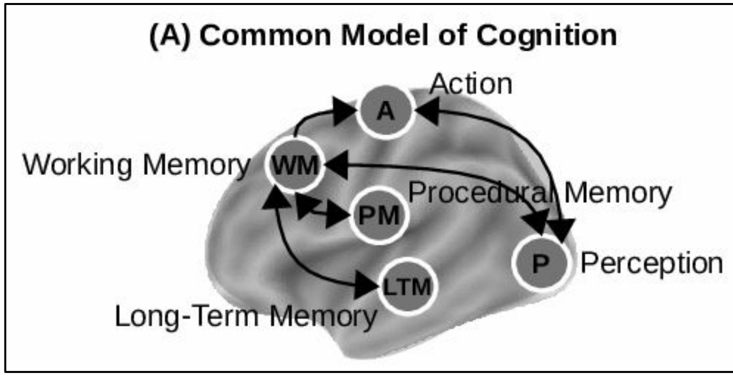


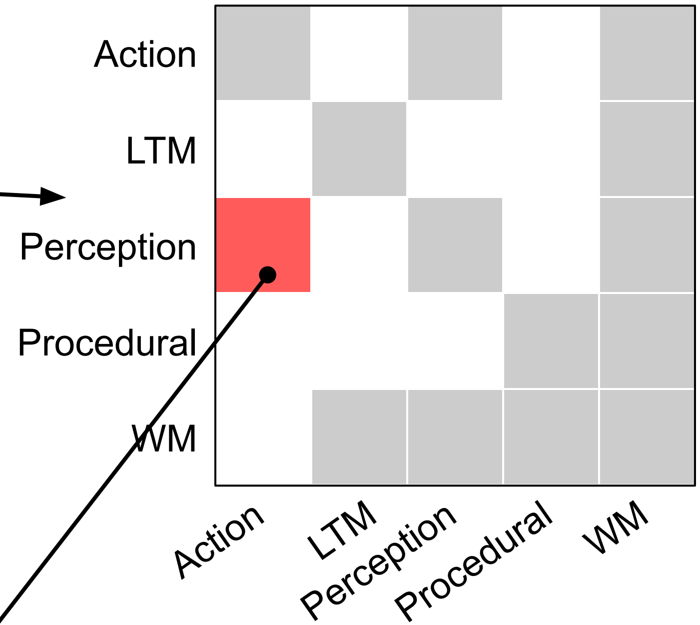
Are All Connections Necessary?

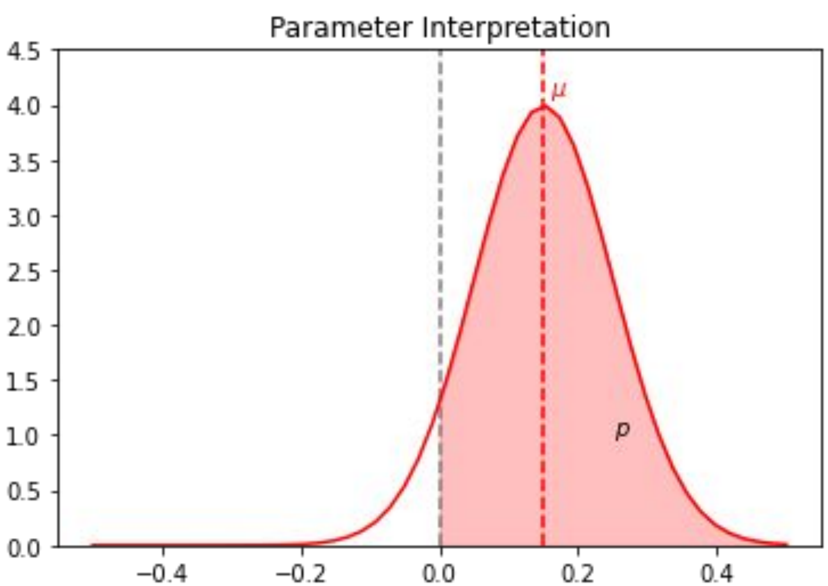




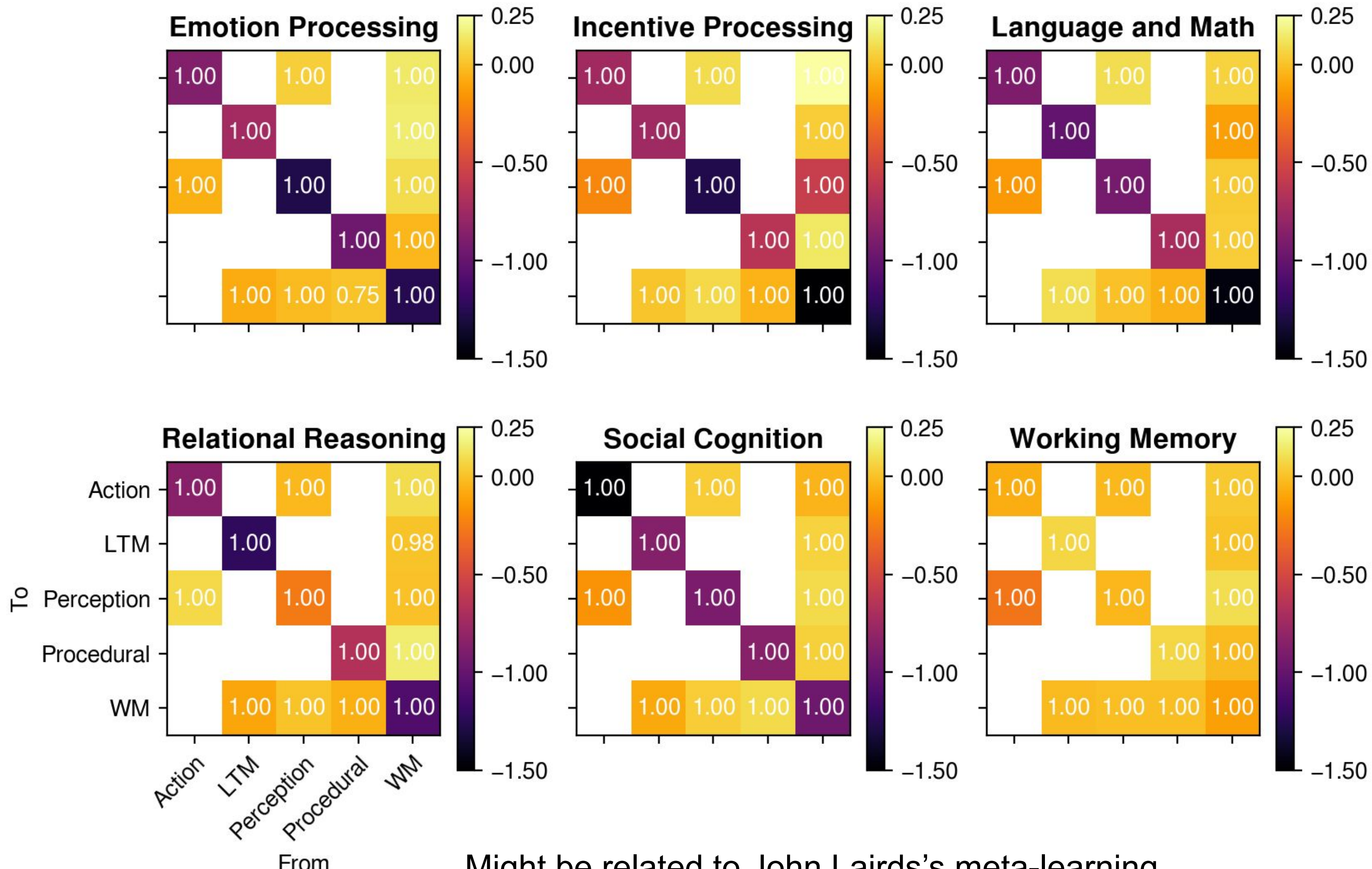
Are All Connections Necessary?







Connectivity Values Across Tasks



Might be related to John Lairds's meta-learning data from his first talk ?

Interim Conclusion

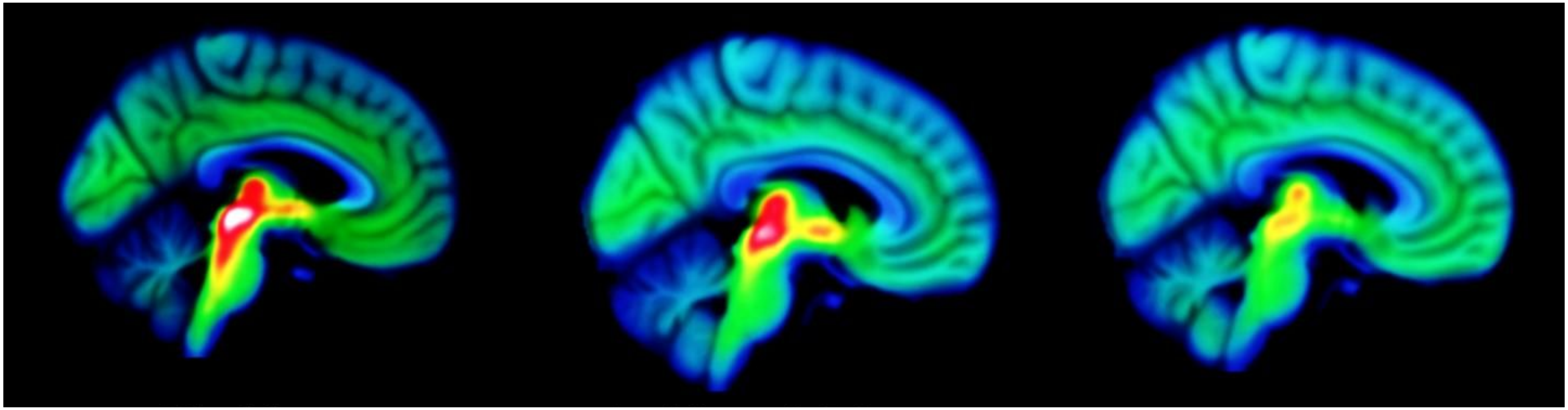
- The CMC works surprisingly well as a high-level architecture for the human brain
  - Fits the **generality** and **superiority** criteria
- The success rests on two factors:
  - A **central hub** (Working Memory/PFC), which might function as a **global workspace**
  - A direct ***P***↔***A*** connection
- Reminescent of dual-control theories
  - Automatic vs. Controlled (Schneider & Shiffrin)
  - Contention Scheduling + Executive (Norman & Shallice)
  - System 1 / System 2 (Kahneman)

Progress Towards Goals (or New Goals)

1. Develop a **methodology** to translate cognitive architectures into testable network brain architectures
   - Language for abstract architecture definition
   - Pipeline to map components to brain and fit models
2. Test and compare **alternative architectures**
   - Task-based fMRI data
   - Resting-state data
3. **Validation** of the results
   - Parameter analysis
   - Alternative methods
   - MEG/EEG data
   - Ground Truth with Clinical Population

Parkinson's Disease (PD)

Progressive Depletion of Dopamine in PD

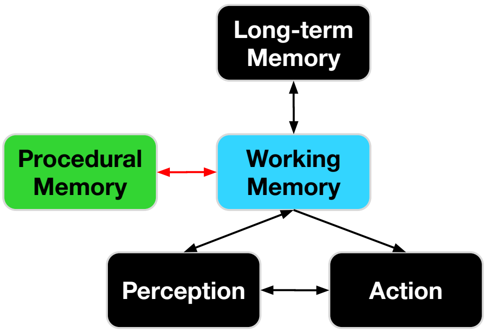


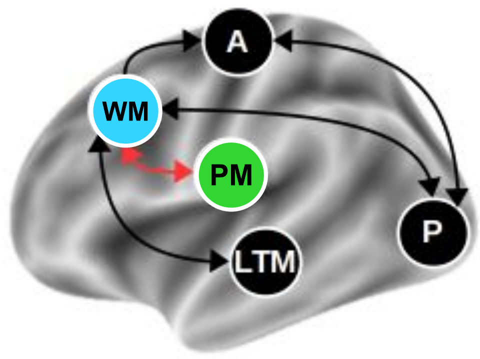
Healthy Controls | PD, Early Stages No Symptoms | PD, Late Stages w/ Symptoms

Why?

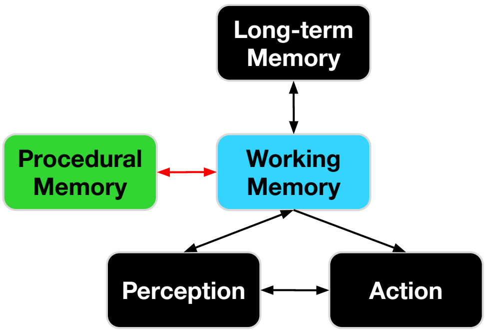
- Main reason: It's the ultimate test
  - Solid **ground truth**
  - Known **etiology** = known localized source
  - known **symptoms** = predictable effects on network parameters
- In addition, opportunity to
  - Examine resting-state data
  - Examine bilateral regions

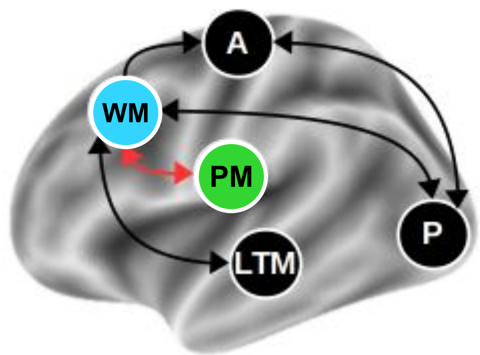
Procedural Memory in the CMC

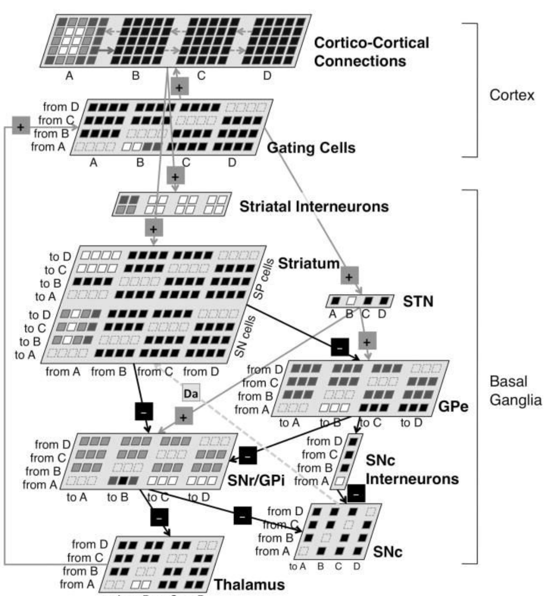




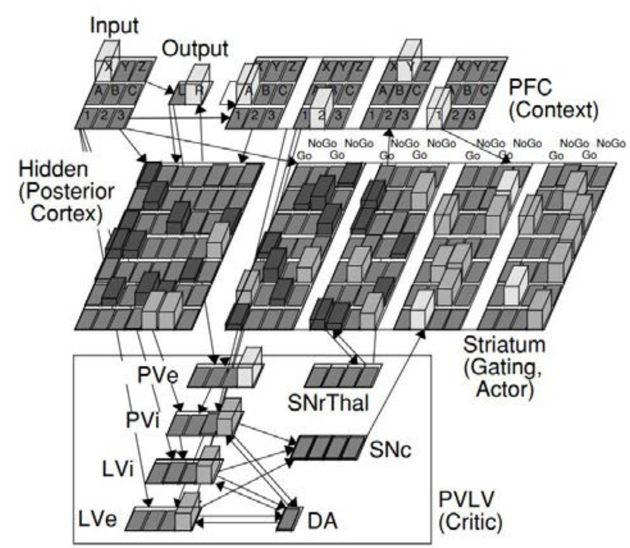
Procedural Memory in the CMC





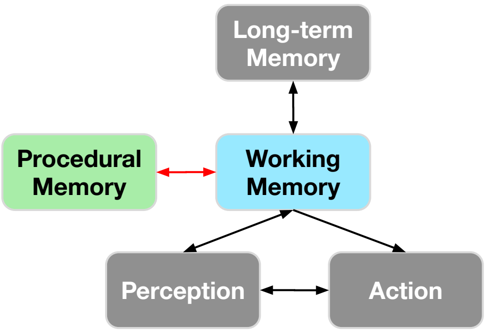


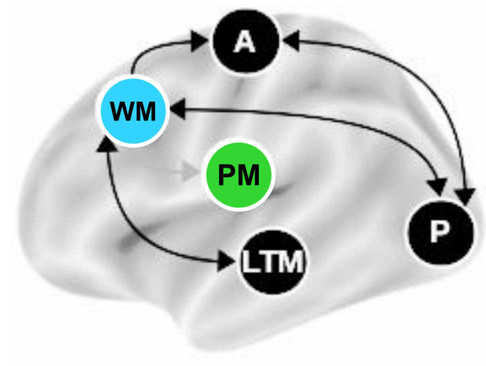
Stocco et al., 2010



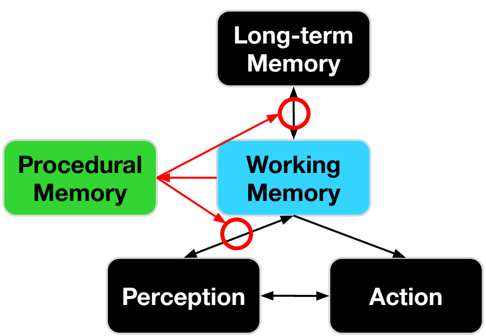
Frank & O'Reilly, 2006

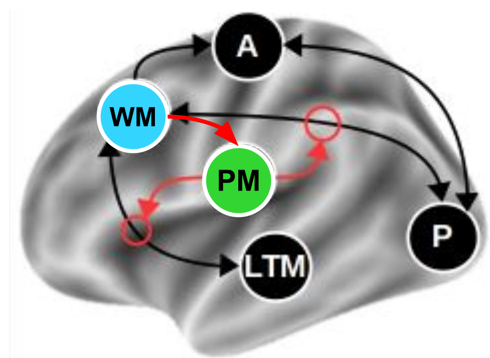
Procedural Memory in the CMC





Direct

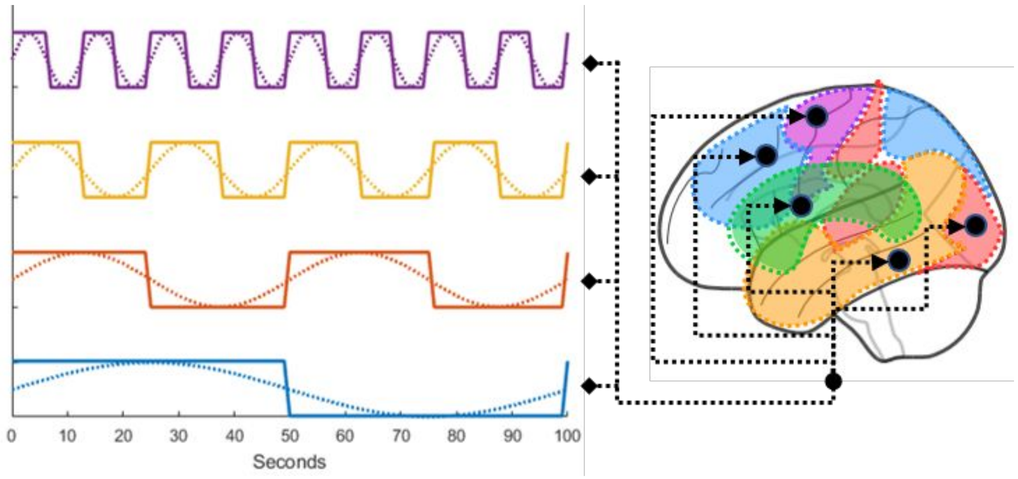




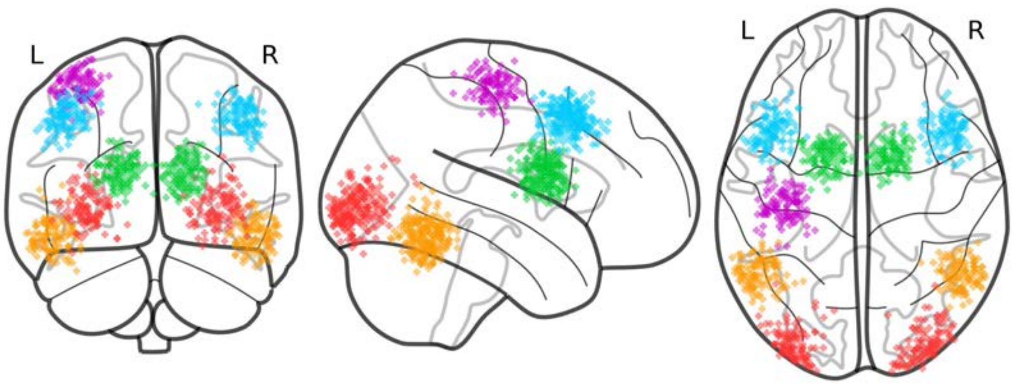
Modulatory

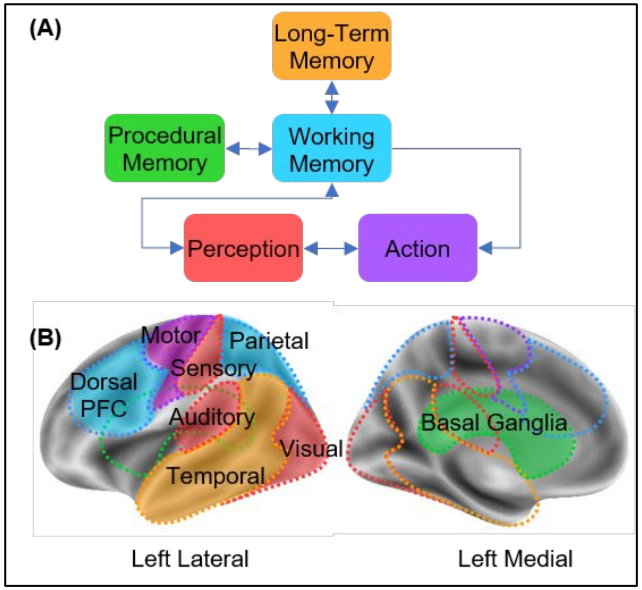
Resting State + Bilateral ROIs

Low-Frequency Regressors Sine Regressors for Resting State

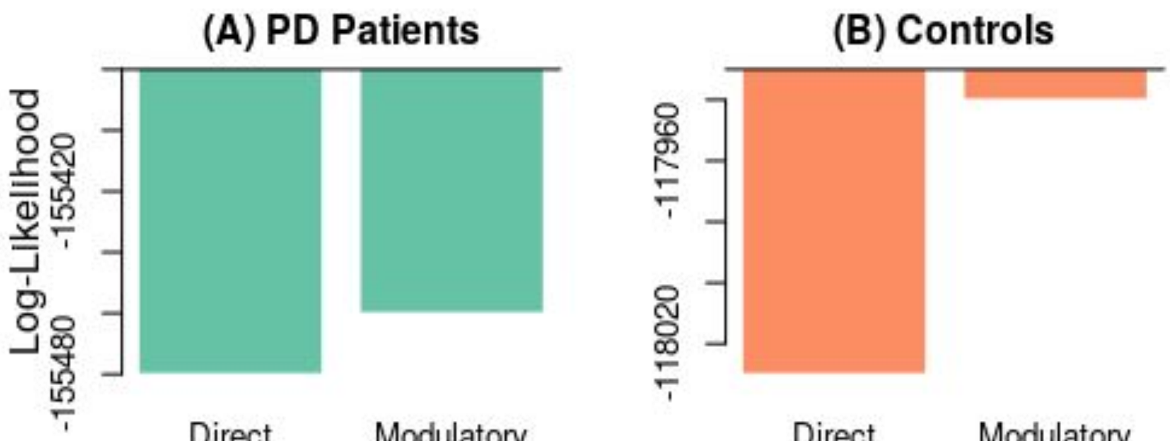


Individualized locations of bilateral ROIs across participants



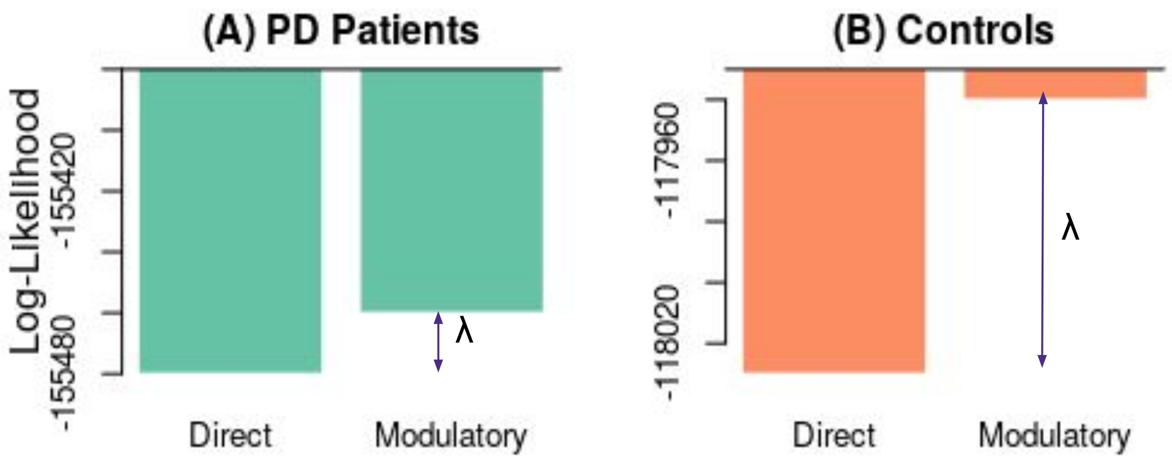


Comparison

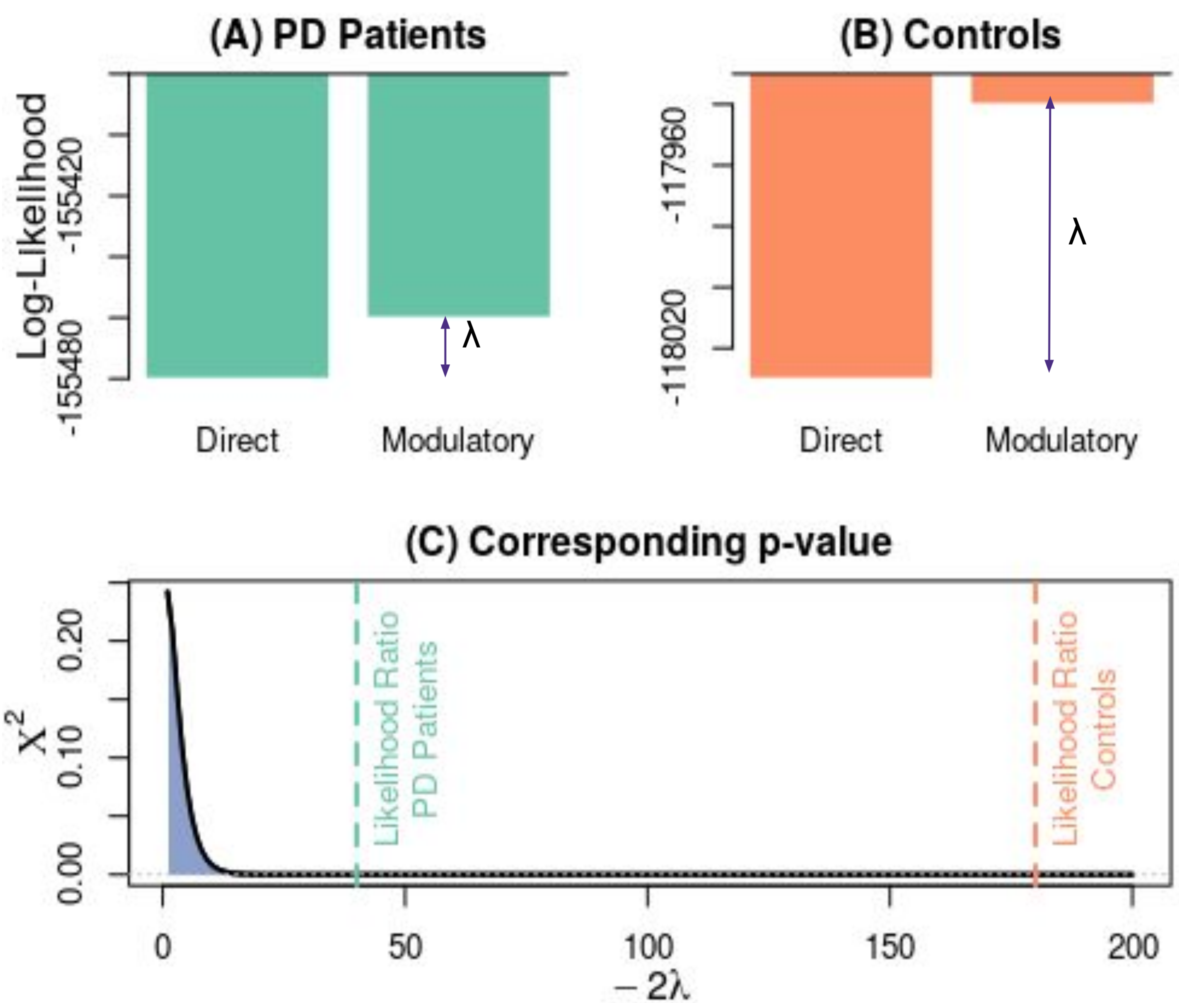


(A) PD Patients
Log-Likelihood
-155420
-155480
Direct
Modulatory
(B) Controls
-117960
-118020
Direct
Modulatory

Comparison


(A) PD Patients
Log-Likelihood
-155420
-155480
λ
Direct
Modulatory
(B) Controls
-117960
-118020
λ
Direct
Modulatory

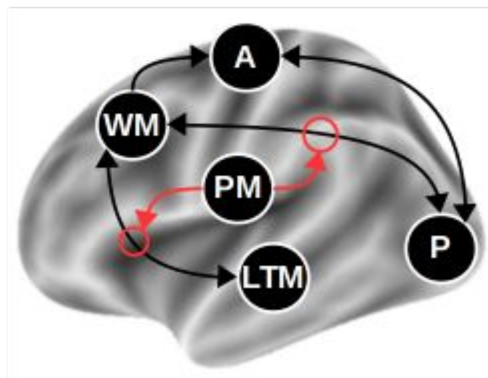
Comparison



(A) PD Patients
Log-Likelihood
-155420
-155480
λ
Direct
Modulatory
(B) Controls
-117960
-118020
λ
Direct
Modulatory
(C) Corresponding p-value
X^2
0.00
0.10
0.20
Likelihood Ratio
PD Patients
Likelihood Ratio
Controls
0
50
100
150
200
-2λ

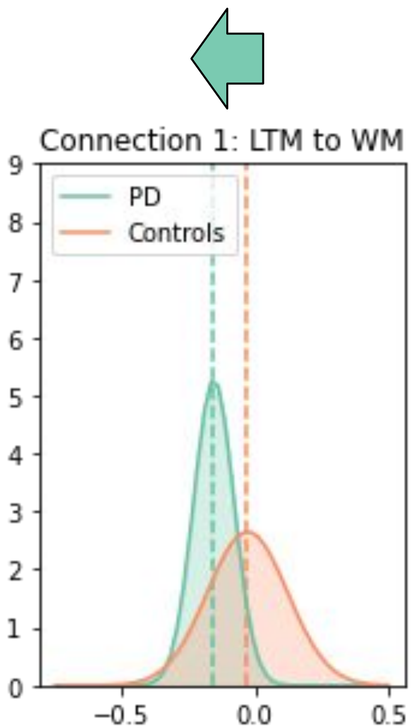
PD-Related Differences in Parameter Values

Modulatory CMC

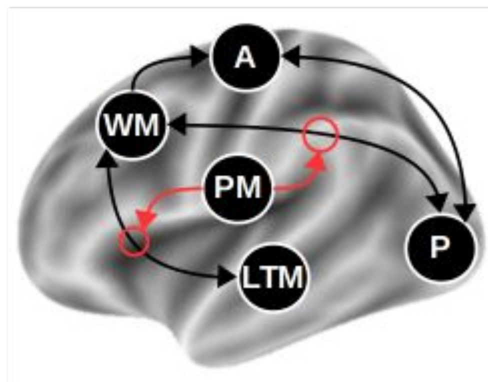


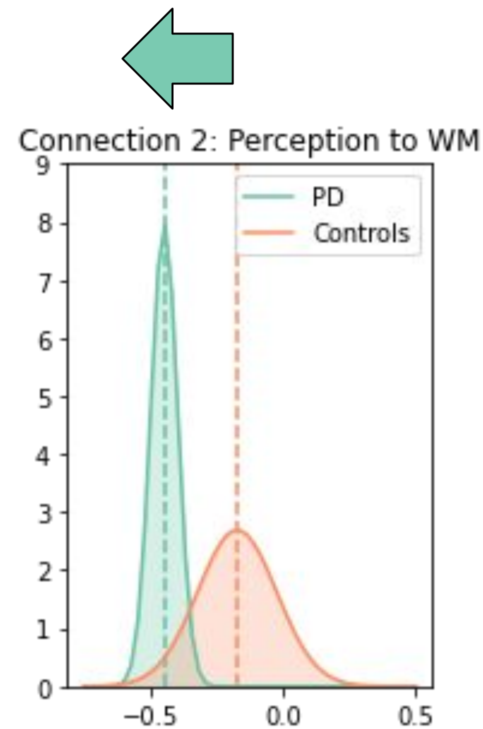
PD-Related Differences in Parameter Values

Modulatory CMC



$p < 0.0001$





$p < 0.0001$

List of Publications, Awards, Honors, etc. Attributed to the Grant

- Stocco, A., Sibert, C., Steine-Hanson, Z., Koh, N., Laird, J., Lebiere, C., & Rosenbloom, P. (under review). Analysis of the Human Connectome Data Supports the Notion of A "Common Model of Cognition" for Human and Human-Like Intelligence Across Domains. *NeuroImage*. [https://www.biorxiv.org/content/10.1101/703777v3]
- Zhou, P. Sense, F., van Rijn, H., & Stocco, A. (under review) Reflections of Idiographic Long-Term Memory Characteristics In Resting-State Neuroimaging Data. *Cognition*.
- Ketola, M., Thompson, S., Madhyastha, T., Grabowski, T., Wapstra, N. & Stocco, A. (under review). *Abnormal modulatory basal-ganglia effective connectivity in Parkinsonism*.
- Ketola, M., Thompson, S., Madhyastha, T., Grabowski, T., & Stocco, A. (2020). Applying the Common Model of Cognition to resting-state fMRI leads to the identification of abnormal functional connectivity in Parkinson's Disease. *Proceedings of the 42nd Annual Meeting of the Cognitive Science Society*.
- Smith, B. M., Chiu, M., Yang, Y., Sibert, C., & Stocco, A. (2020) Modeling the effects of post-traumatic stress on hippocampal volume. *Proceedings of the 18th International Conference on Cognitive Modeling.* **[Winner of the Allen Newell prize for best student paper]**