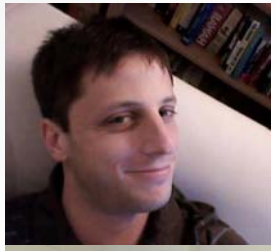
Probabilistic Models for Brain Data Analysis

Sanmi Koyejo
CS & Beckman, University of Illinois

Outline

- **Part 1:** Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition
- **Part 2:** A Hierarchical Model for Time-Varying Functional Connectivity

Part 1:
Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition



Timothy Rubin @ Indiana => SurveyMonkey

Michael Jones @ Indiana

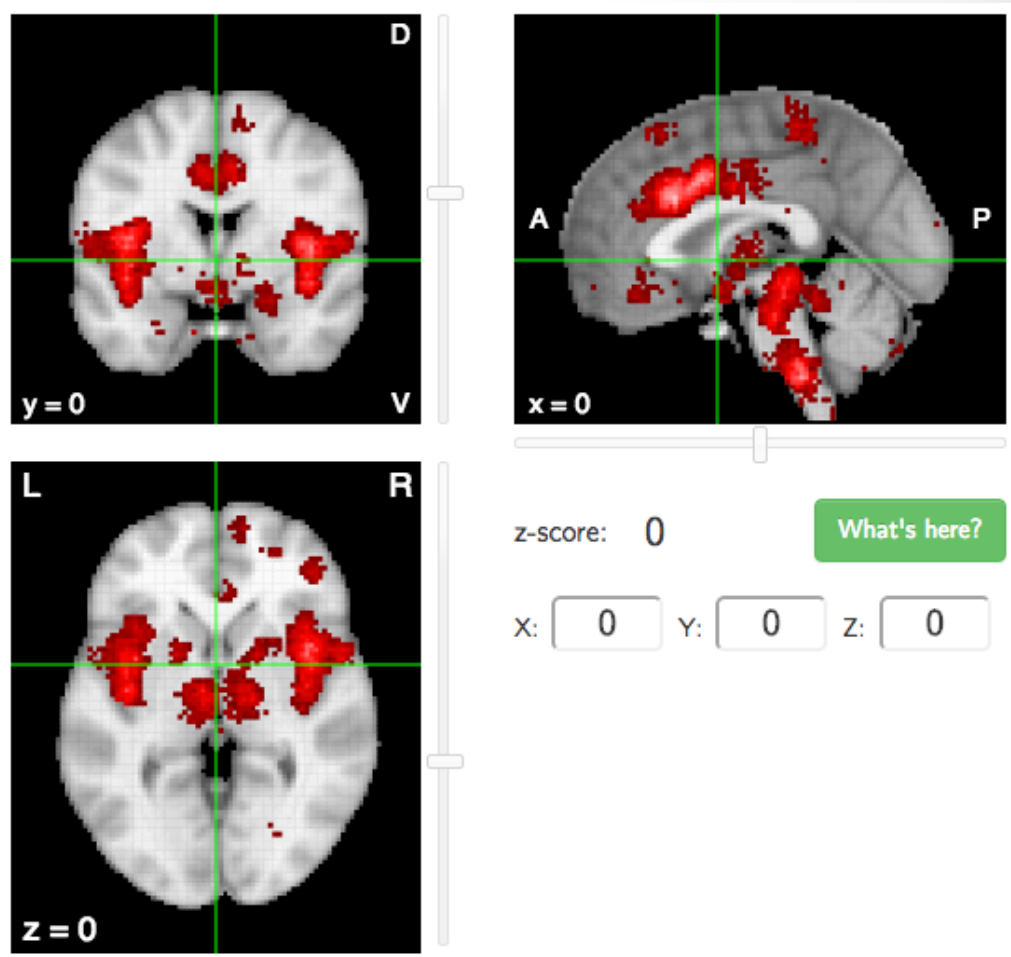
Chris Gorgolewski @ Stanford

Russ Poldrack @ Stanford

Tal Yarkoni @ UT Austin

Encoding

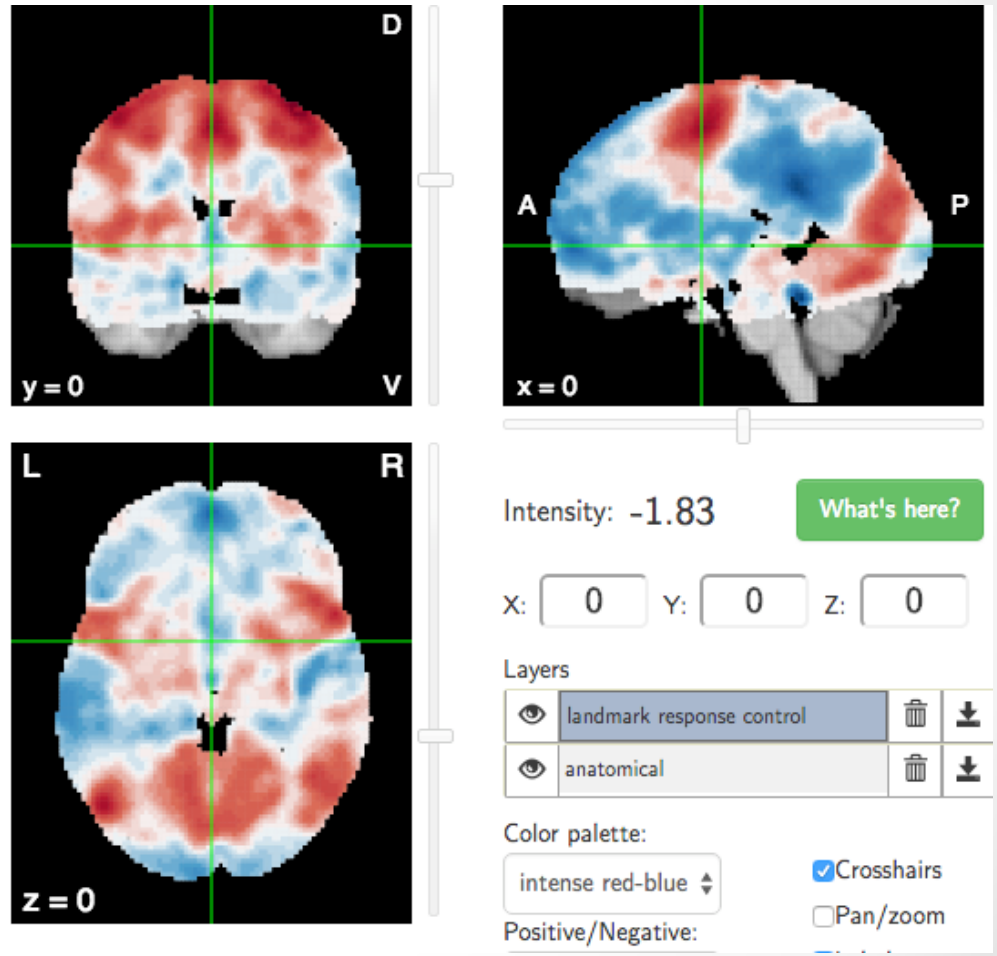
Pain



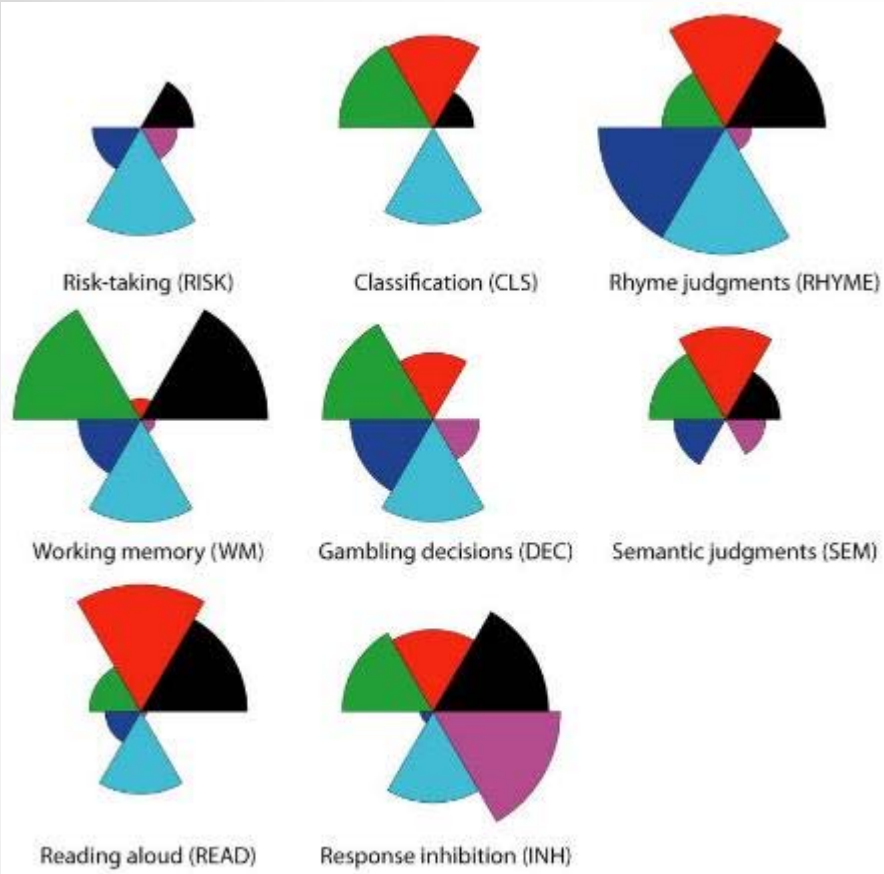
Source: Neurosynth

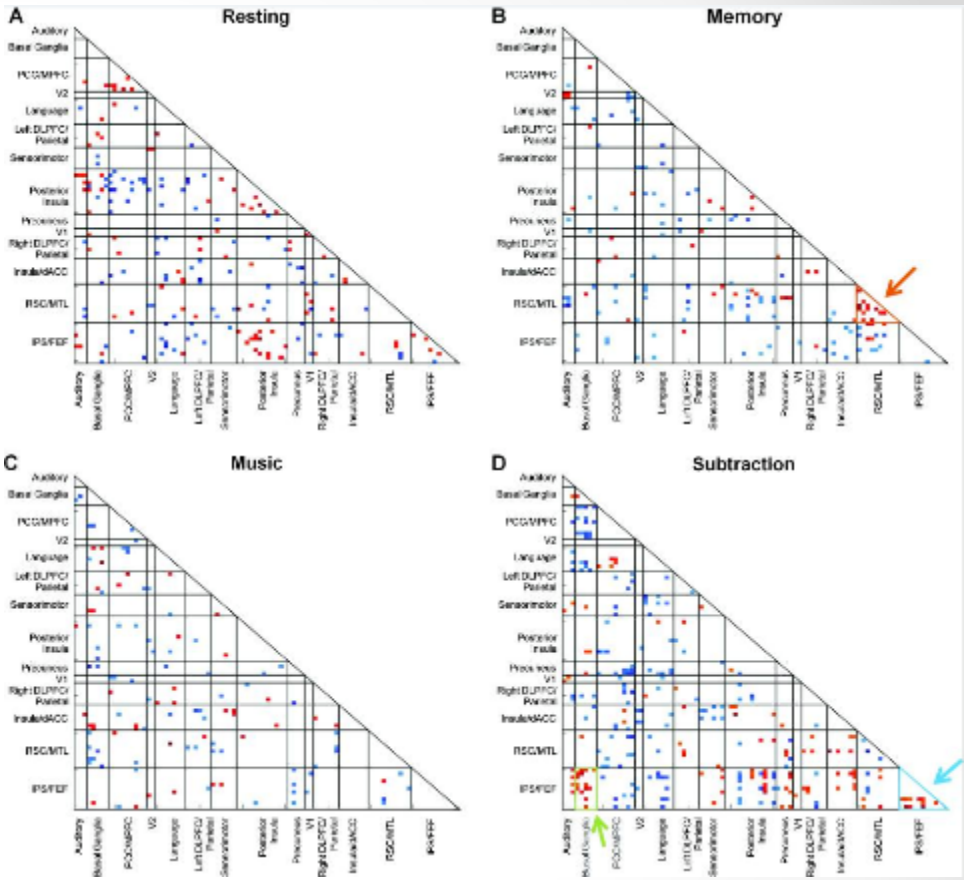
Decoding

premotor	0.419
motor	0.365
movements	0.364
parietal	0.359
visual	0.346
premotor cortex	0.334
execution	0.326
intraparietal	0.325
finger	0.322
hand	0.317



Source: Neurosynth





Classifier loadings for 8 tasks (Poldrack et. al. 2009)

84% accuracy across subjects, 64% across tasks

Distinct connectivity patterns for 4 cognitive states (Shirer et. al. 2008)

85% LOOCV accuracy

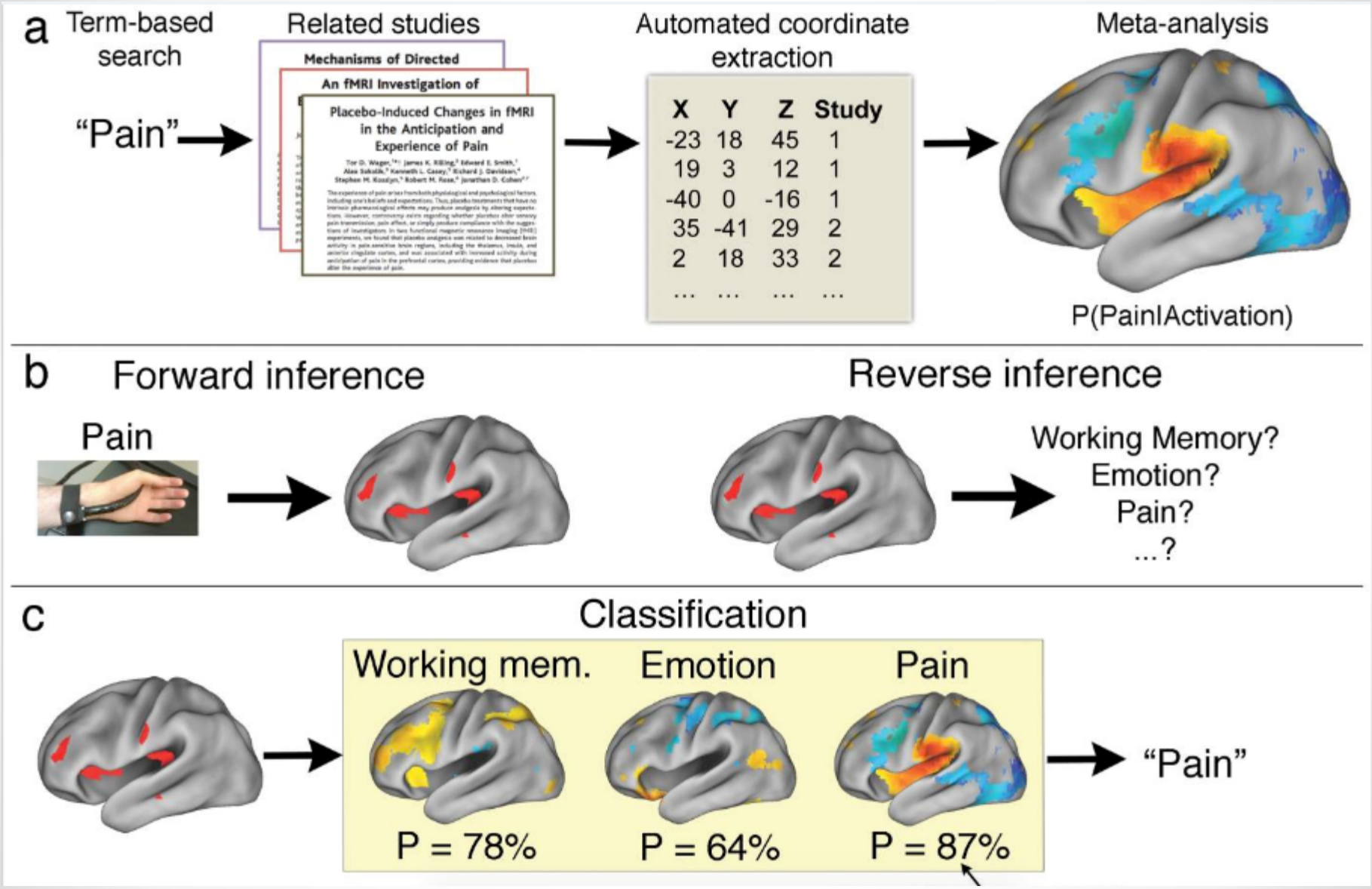
Open ended decoding via meta-analysis

- Brainmap (Fox et. al. 1994)

brainmap.org

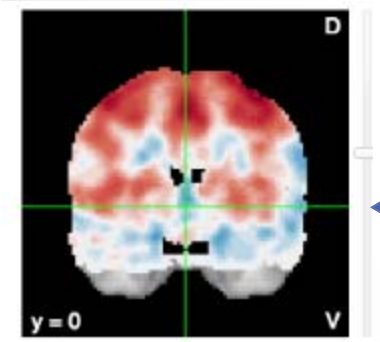
- Neurosynth (Yarkoni et. al. 2011)

neurosynth.org

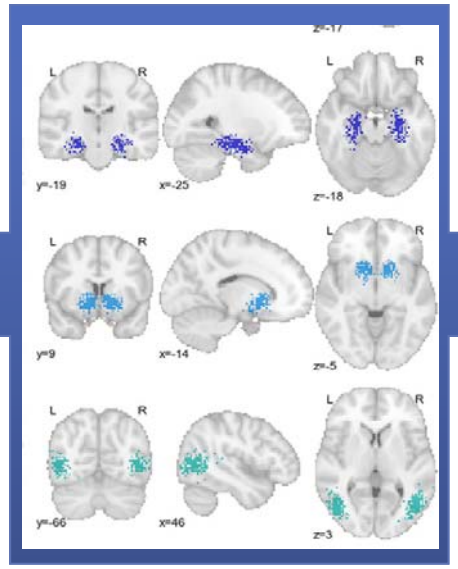


Neurosynth (Yarkoni et. al. 2011)

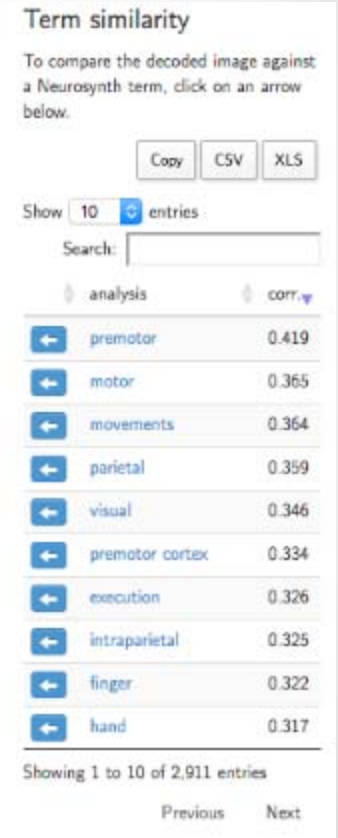
Open ended Encoding & Decoding



Brain Image



Latent Representation



Cognitive Terms

Straightforward decoding of whole brain maps

Predicting risky choices from brain activity patterns

Sarah M. Helfinstein[a,1], Tom Schonberg[a], Eliza Congdon[b,c], Katherine H. Karlsgodt[d,e], Jeanette A. Mumford[f], Fred W. Sabb[b,c,g], Tyrone D. Cannon[h,i], Edythe D. London[b,c,g,j], Robert M. Bilder[b,c,g,k], and Russell A. Poldrack[a,f,l]

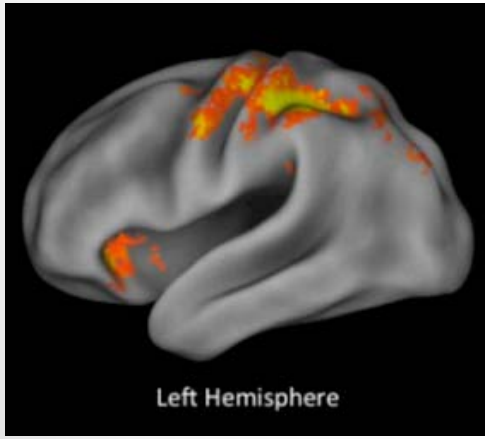


Table 2. Pearson correlations between searchlight classification map and NeuroSynth term-based reverse inference activation maps

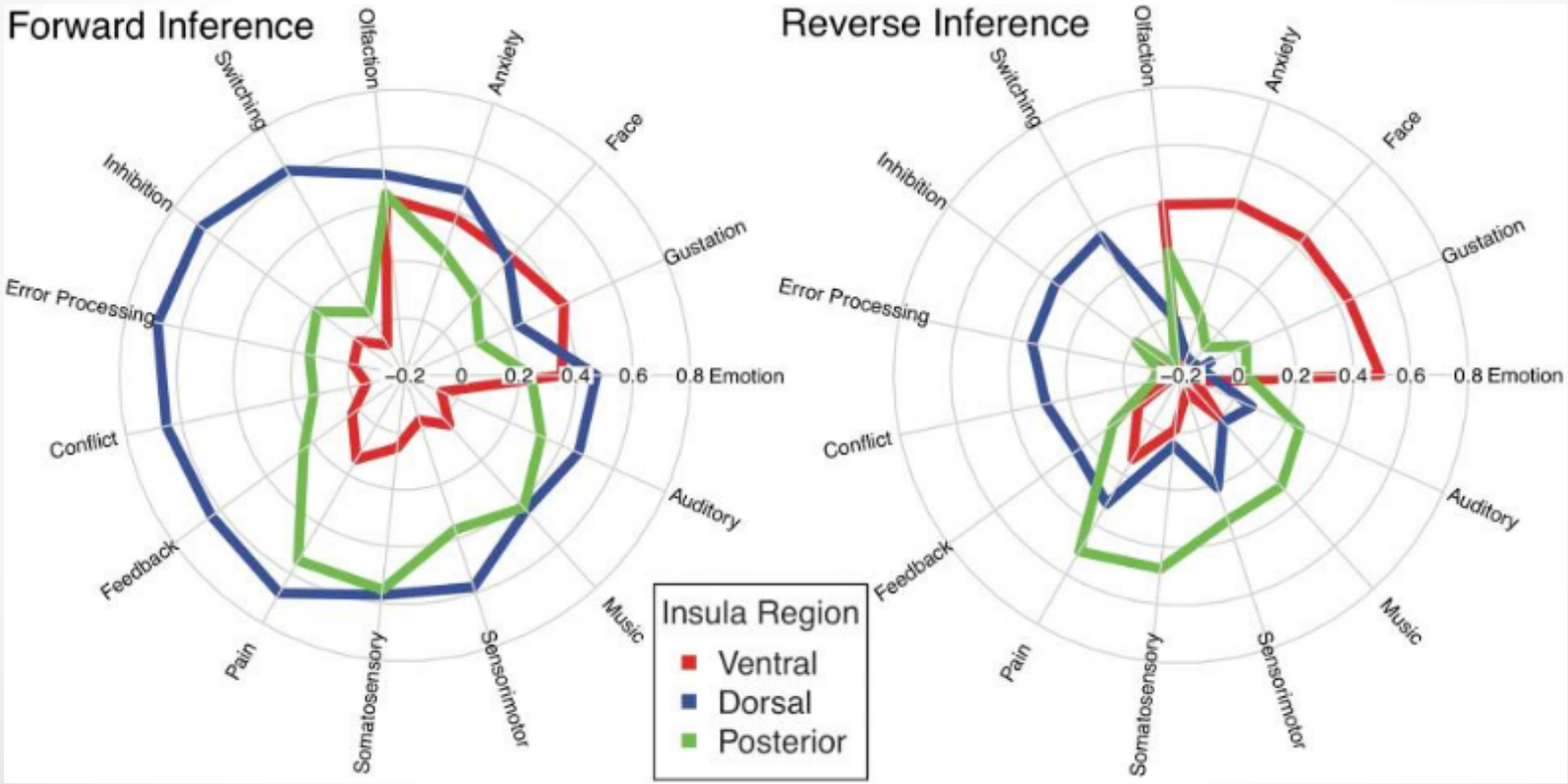
Term	Correlation (*r*)
Control	0.1451
Working	0.1159
Numerical	0.1157
Letter	0.1081
Attention	0.1062
Correct	0.1060
Cue	0.0995
Preparatory	0.0970
Load	0.0959
Hand	0.0924

The 10 most highly correlated terms are listed.

Helfinstein et. al. (2014)

Decoding the Role of the Insula in Human Cognition: Functional Parcellation and Large-Scale Reverse Inference

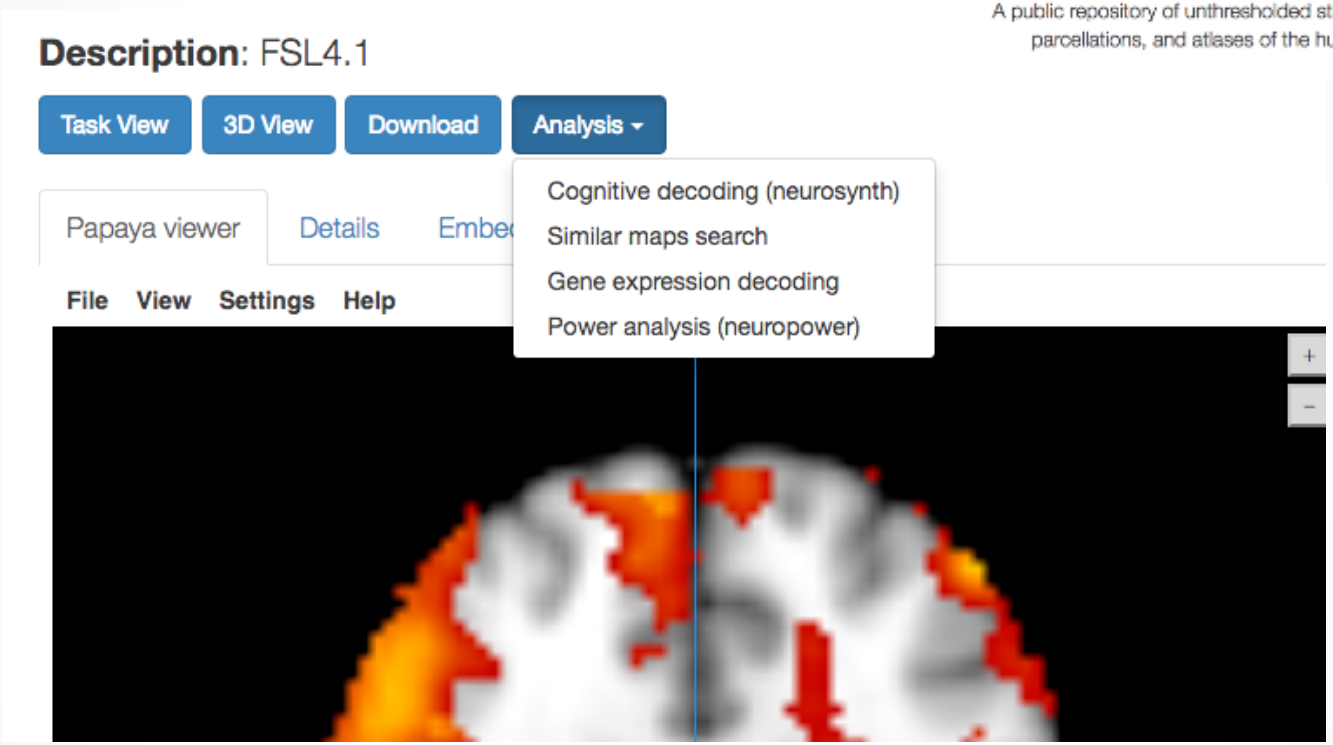
Luke J. Chang[1,2], Tal Yarkoni[3], Mel Win Khaw[4] and Alan G. Sanfey[1,5,6]



Chang et. al. 2012

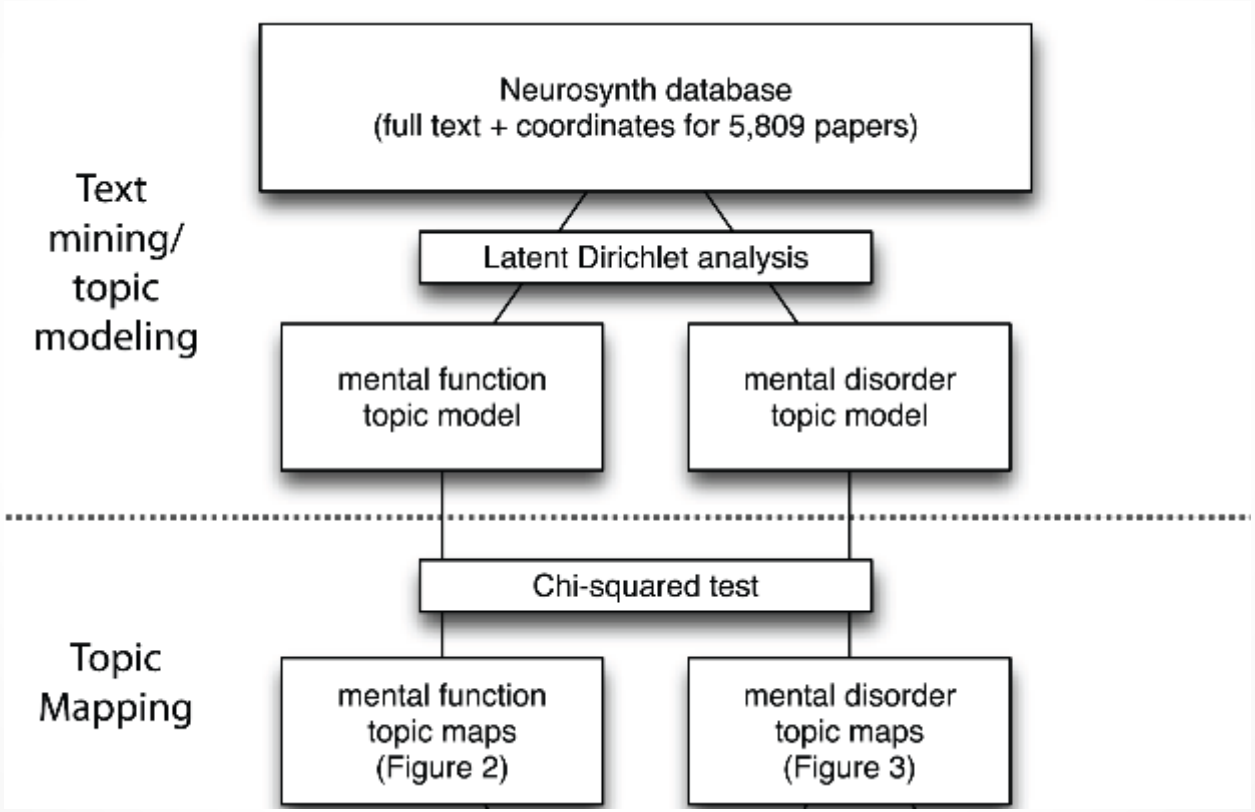
NEUROVAULT

A public repository of unthresholded statistical maps, parcellations, and atlases of the human brain



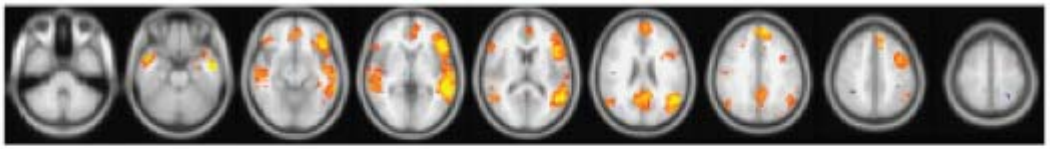
Gorgolewski et. al. (2015)

Terms are fine grained

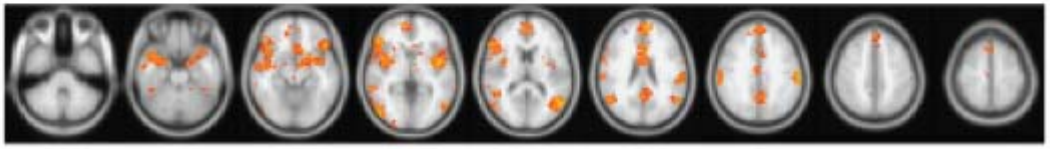


Poldrack et. al. (2012)

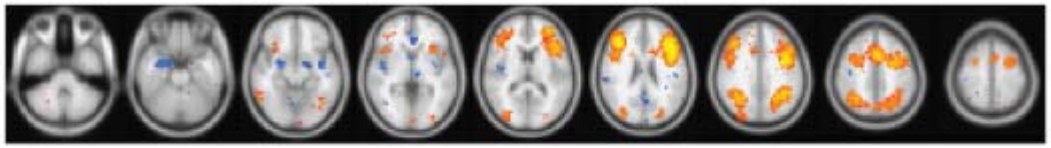
Topic 71 (115 docs):
narrative, discourse, comprehension, memory, discourse_processing



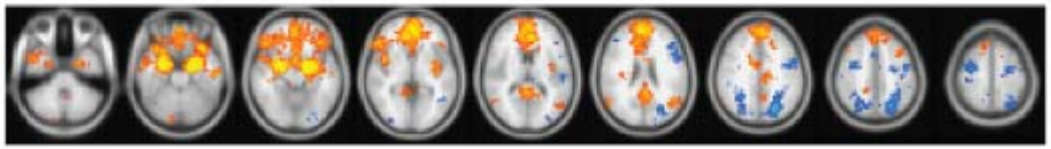
Topic 108 (128 docs):
empathy, pain, theory_of_mind, awareness, facial_expression



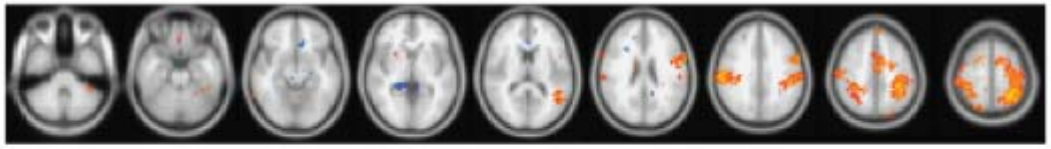
Topic 61 (441 docs):
memory, working_memory, maintenance, visual_working_memory, spatial_working_memory



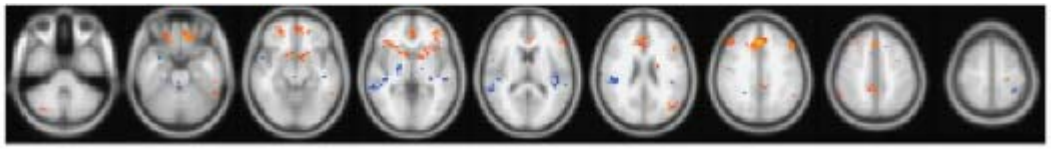
Topic 93 (495 docs):
emotion, valence, arousal, attention, focus



Topic 20 (497 docs):
action, movement, goal, context, perception



Topic 86 (519 docs):
decision, decision_making, choice, fixation, uncertainty



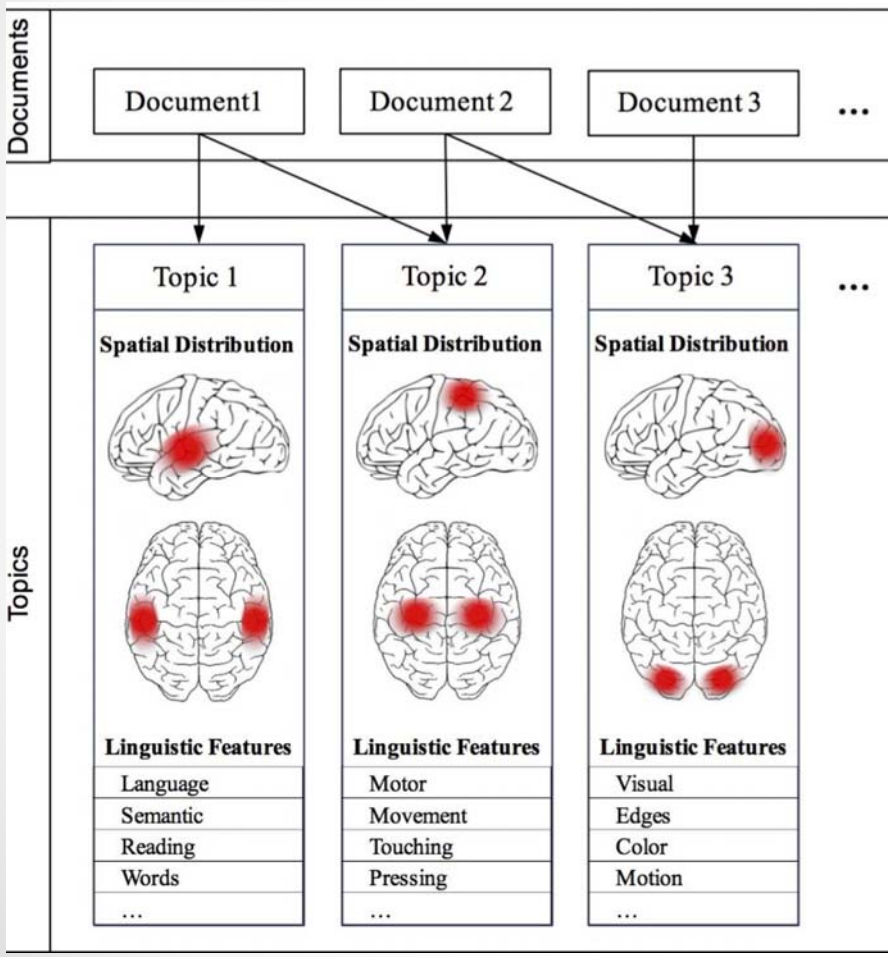
Poldrack et. al. (2012)

Open Issues

- Current decoding approaches lack a formal model
- Same issue when combined with topic mapping
- Does not attempt to identify any latent structure that presumably makes such mappings useful for example, individual brain regions or functional brain networks that correspond to specific cognitive processes

Generalized Correspondence LDA (GC-LDA)

Generalized Correspondence-LDA

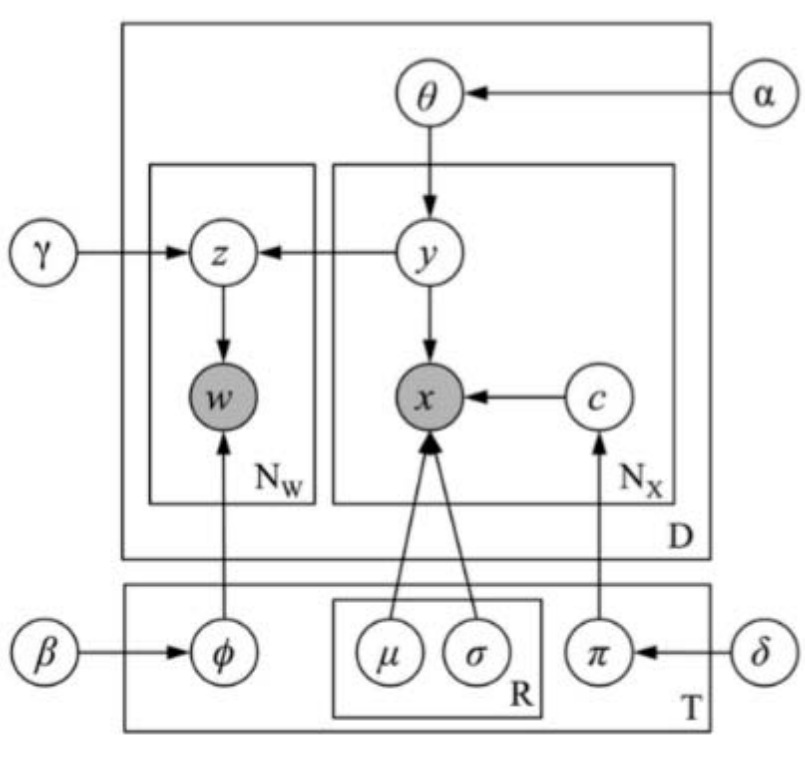


- A generalization of correspondence-LDA (Blei, Ng, & Jordan, 2003)
- Each latent topic is associated with both a spatial distribution and a set of word tokens
- Explicitly designed to produce *region-like* topics
- We apply a symmetric constraint to produce bilaterally distributed topics

Some useful properties

- Facilitates the learning of interpretable latent structures from a mass of superficial braincognition mappings
- Provides the ability to decode bi-directionally
- Enables principled generation of novel exemplars or combinations of events that have never been seen before (e.g., what pattern of brain activity would a task combining painful stimulation and phonological awareness produce?).

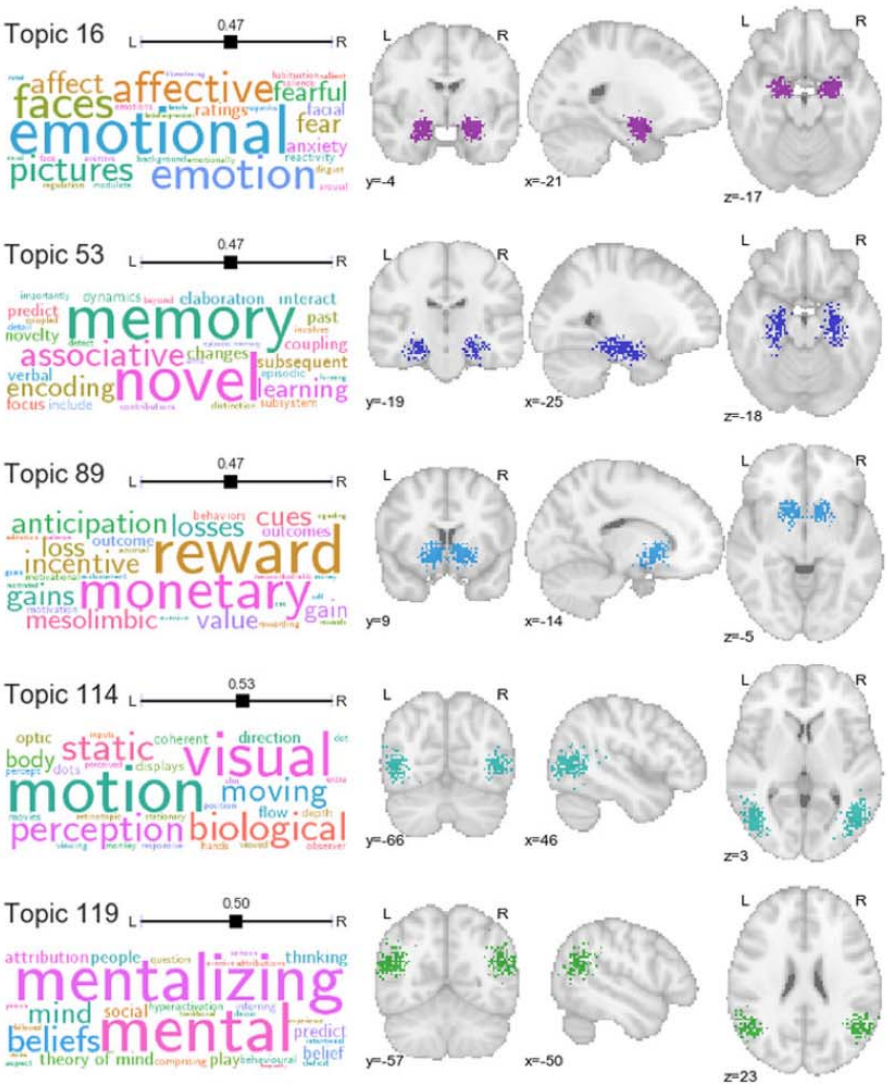
Generative Model



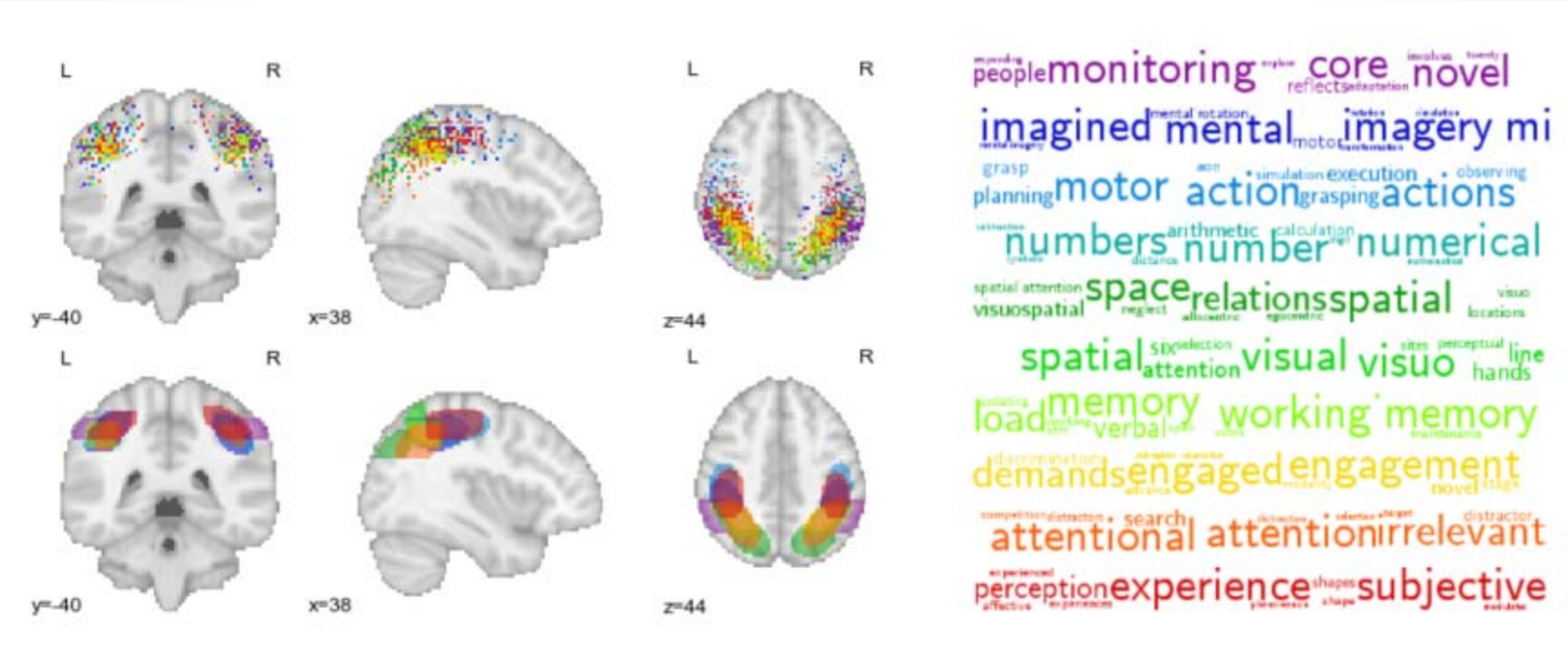
- Each document is a collection of spatial activations and words
- Each "topic" is a combination of a linguistic topic and a spatial activation topic
- Model is trained using MCMC sampling
- Parameter estimation via maximum likelihood with bi-lateral constraints

Neurosynth database

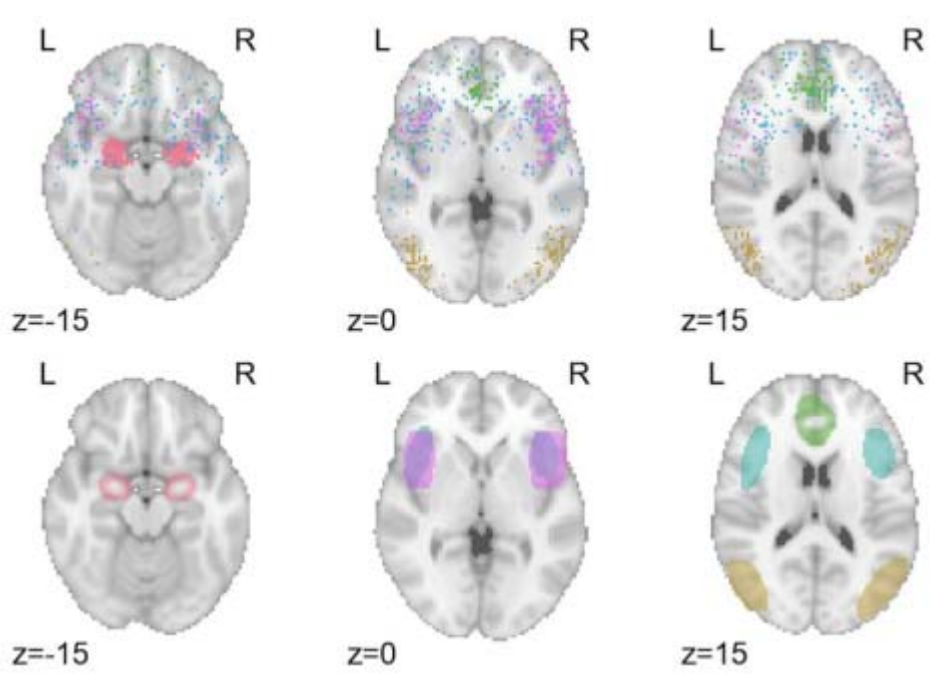
- 11,362 total publications
- Avg. 35 peak activation tokens per document
- Avg. 46 word tokens after preprocessing
- Approximately 400k activation and 520k word tokens in total
- Learned model with 200 topics


Topic 16
0.47
affect affective fearful faces ratings facial fear emotional anxiety pictures emotion
y=-4 x=-21 z=-17
Topic 53
0.47
importantly dynamics elaboration interact predict memory past novelty coupling associative changes subsequent verbal encoding novel learning focus include distinction subsystem
y=-19 x=-25 z=-18
Topic 89
0.47
anticipation losses cues behavior outcome outcomes loss incentive reward gains monetary gain mesolimbic value
y=9 x=-14 z=-5
Topic 114
0.53
optic static coherent direction body displays visual motion moving flow perception biological
y=-66 x=46 z=3
Topic 119
0.50
attribution people question thinking mentalizing mind social hyperactivation beliefs mental predict belief theory of mind comprising play
y=-57 x=-50 z=23
Topic 130
0.50
intensity chronic hyperalgesia nociceptive perceived ratings thermal neuropathic mechanical somatosensory visceral cold matrix heat sensory allodynia painful noxious stimulation
y=-4 x=39 z=11
Topic 111
0.39
switch flexibility eye direction target antisaccade switching saccades saccade task set
y=-4 x=-34 z=48
Topic 96
0.51
engaged question people person focused scale social mental view mind
y=0 x=54 z=-23
Topic 98
0.48
outcomes contexts subsequent ratings predict values value outcome preferences valuation subjective probability preference reflected contextual amount
y=46 x=-1 z=-5
Topic 65
0.41
abilities poorer play/learning selection fact total digit verbal free recall number relationships memory relation twenty evaluated completed names weaker addition name
y=31 x=0 z=47

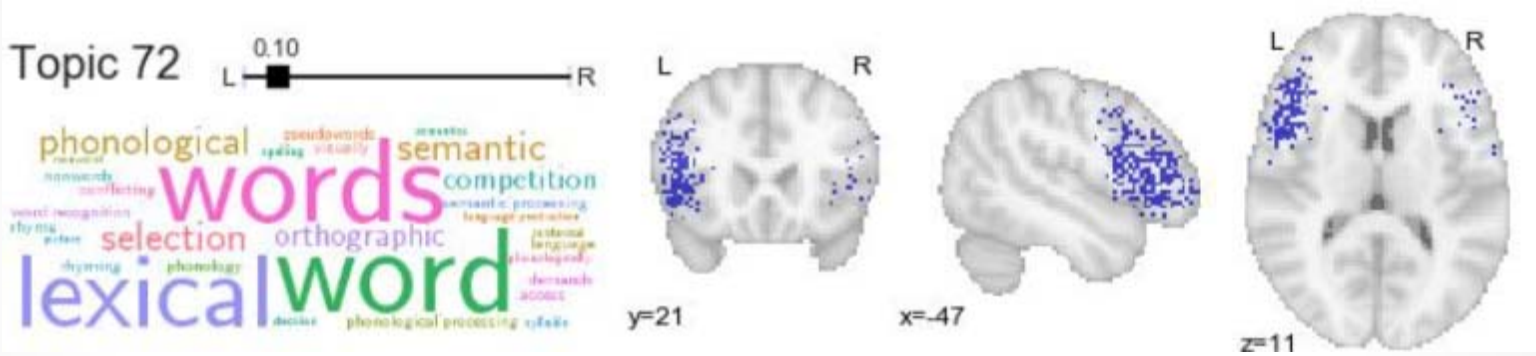
Overlapping topics in Parietal cortex

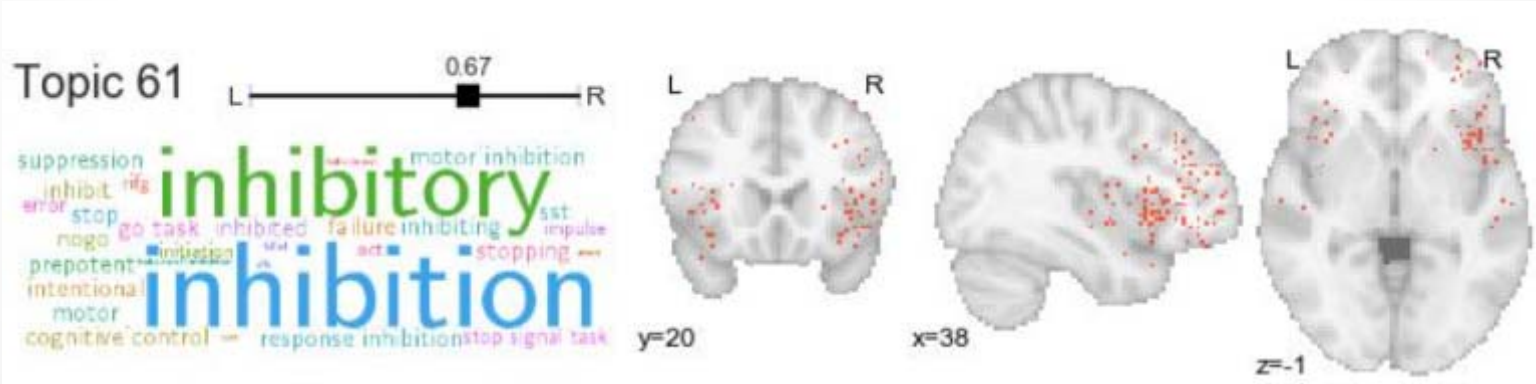


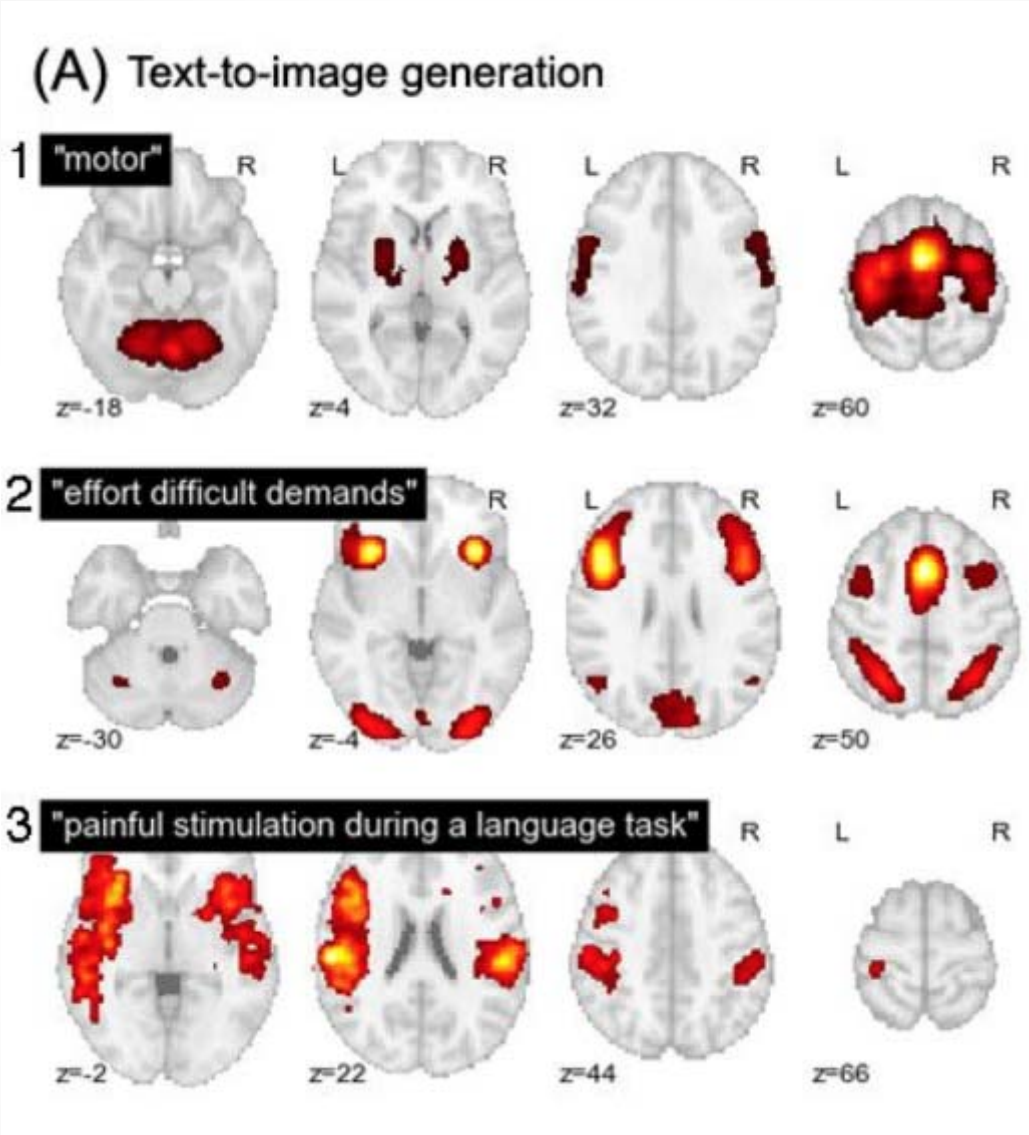
Compositional nature of cognitive states related to emotion

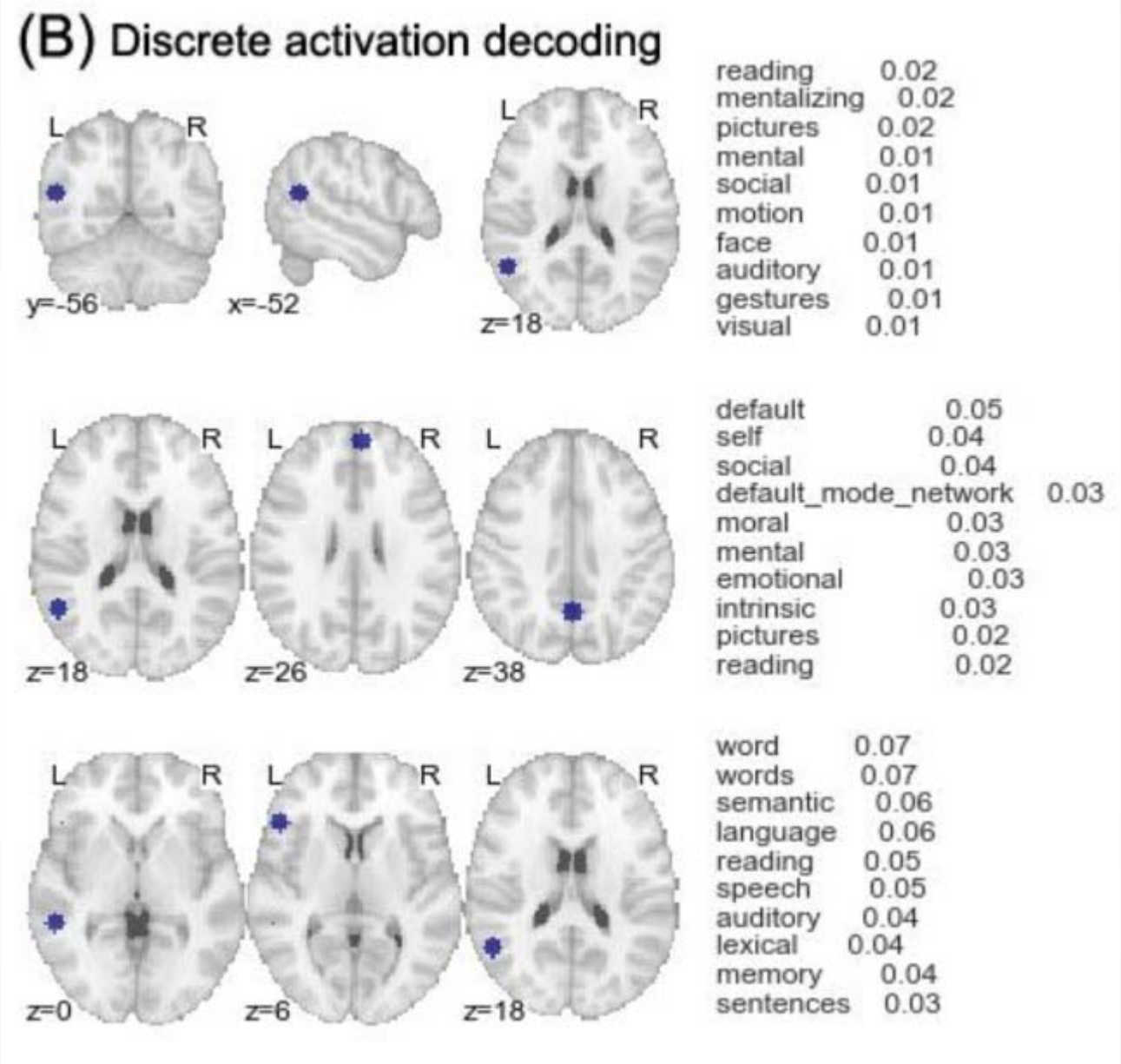


Datadriven window into lateralization of function

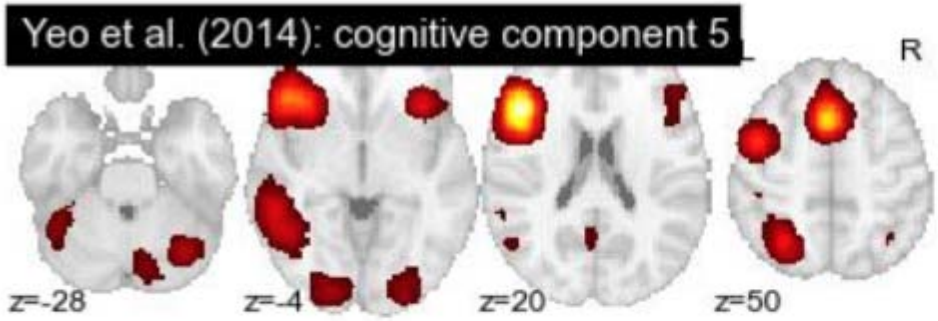




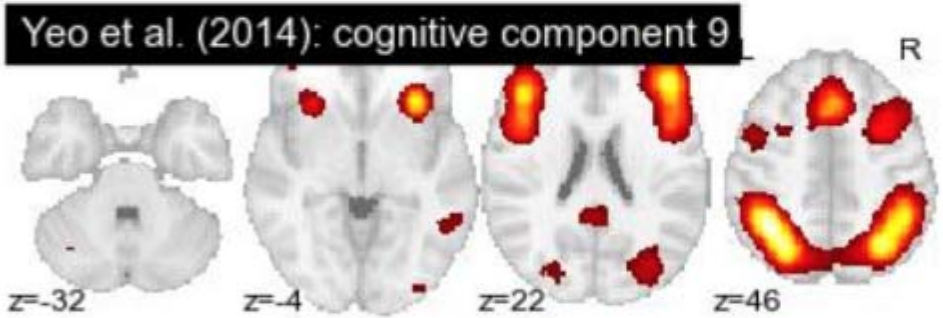

(A) Text-to-image generation
1 "motor"
R
L
R
L
R
L
R
z=-18
z=4
z=32
z=60
2 "effort difficult demands"
R
L
R
L
R
z=-30
z=-4
z=26
z=50
3 "painful stimulation during a language task"
R
L
R
z=-2
z=22
z=44
z=66


(B) Discrete activation decoding
L
R
y=-56
x=-52
z=18
reading 0.02
mentalizing 0.02
pictures 0.02
mental 0.01
social 0.01
motion 0.01
face 0.01
auditory 0.01
gestures 0.01
visual 0.01
z=18
z=26
z=38
default 0.05
self 0.04
social 0.04
default_mode_network 0.03
moral 0.03
mental 0.03
emotional 0.03
intrinsic 0.03
pictures 0.02
reading 0.02
z=0
z=6
z=18
word 0.07
words 0.07
semantic 0.06
language 0.06
reading 0.05
speech 0.05
auditory 0.04
lexical 0.04
memory 0.04
sentences 0.03

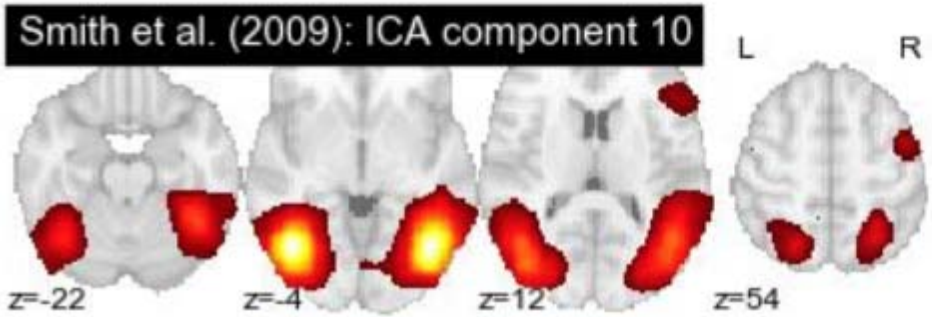
(C) Whole-brain image decoding



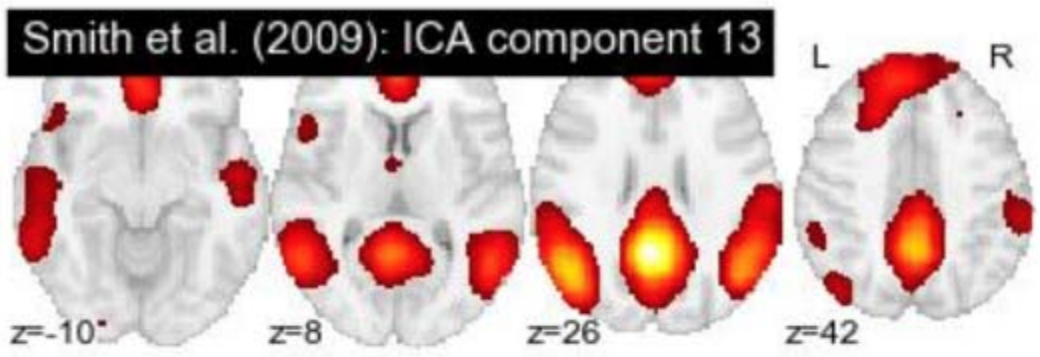
word	0.015
lexical	0.015
words	0.014
phonological	0.013
demands	0.012
syntactic	0.012
sentences	0.011
matching	0.011
manipulation	0.01
sentence	0.01



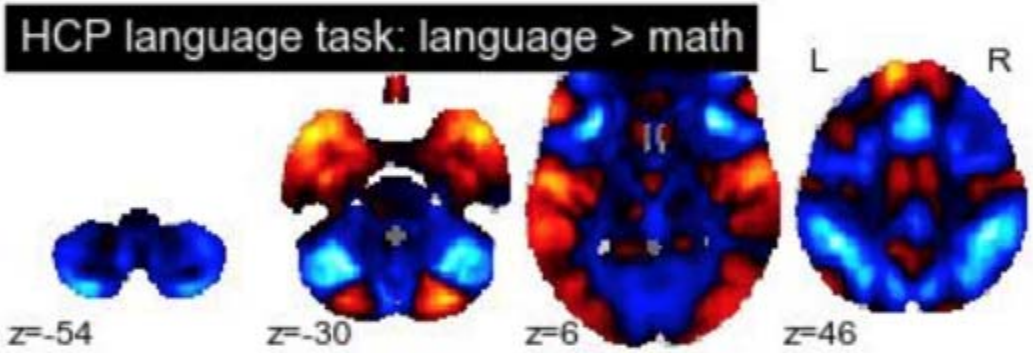
working_memory	0.015
demands	0.011
numerical	0.01
numbers	0.0099
manipulation	0.0096
arithmetic	0.0092
number	0.0085
load	0.0084
verbal	0.0083
maintenance	0.0083



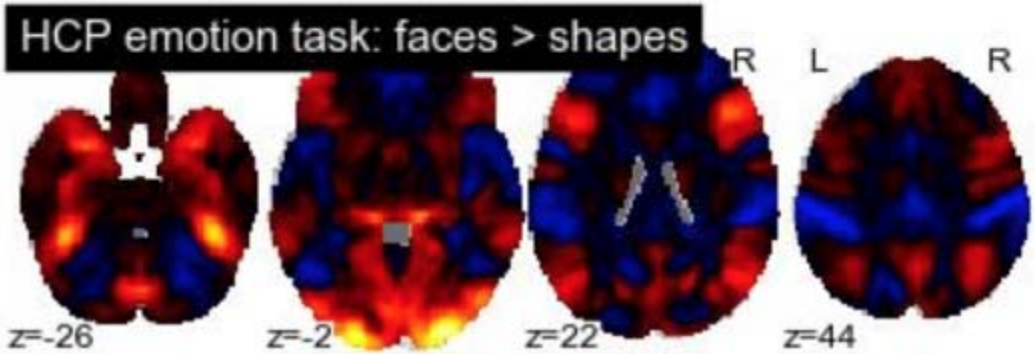
motion	0.059
visual	0.035
face	0.033
faces	0.028
static	0.024
biological	0.021
reading	0.021
moving	0.02
gestures	0.019
pictures	0.018



default	0.024
default_mode_network	0.021
mentalizing	0.014
intrinsic	0.013
episodic	0.012
self	0.011
memories	0.011
seed	0.011
motion	0.011
autobiographical	0.01

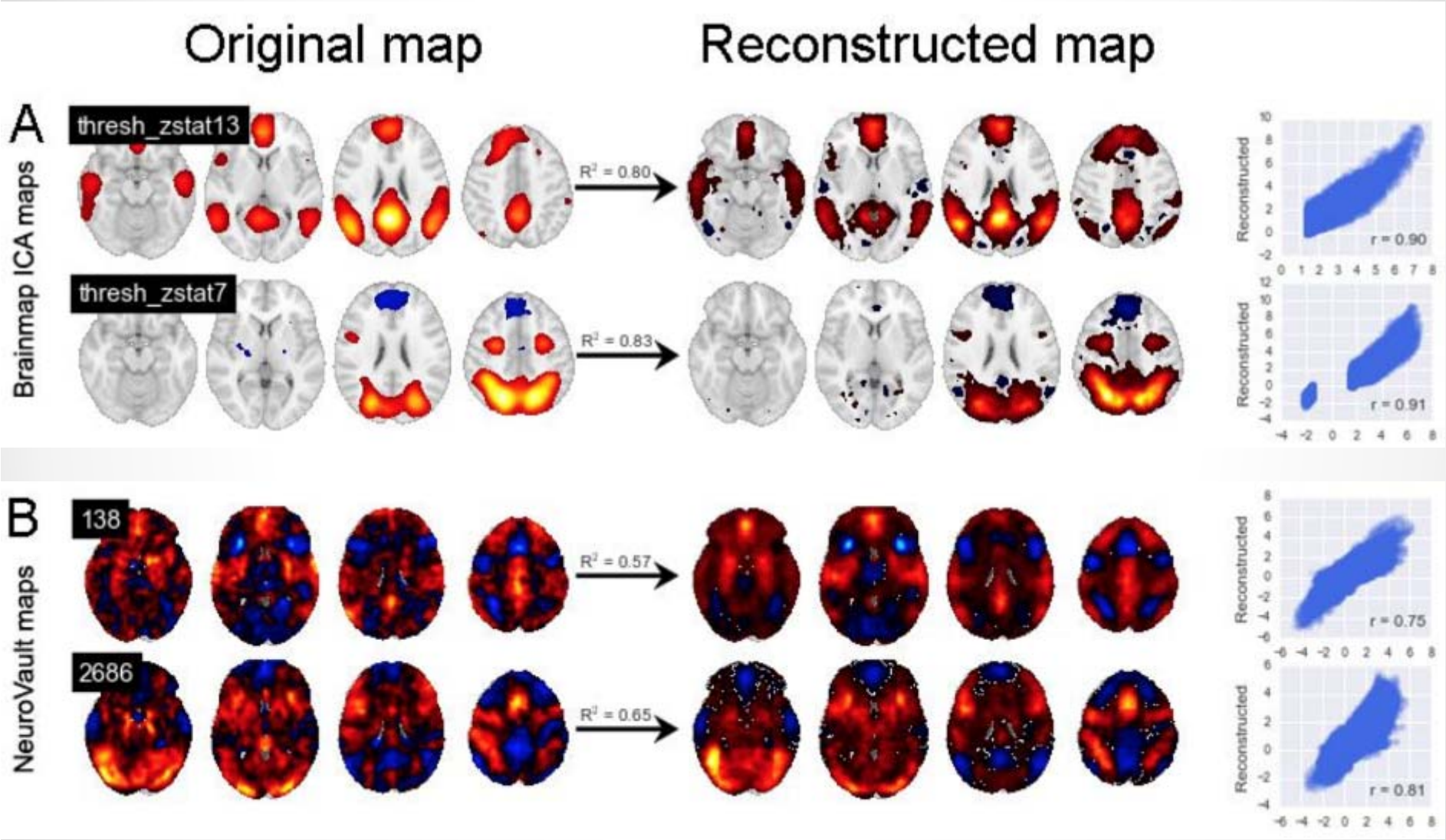


speech	0.062
auditory	0.06
voice	0.055
sounds	0.047
sound	0.03
acoustic	0.029
motion	0.026
default	0.025
listening	0.022
speaker	0.02



faces	0.082
face	0.076
visual	0.071
matching	0.049
facial	0.047
pictures	0.042
disgust	0.041
words	0.04
word	0.04
emotional	0.039

Map Reconstruction



Limitations

- Specificity of the extracted topics is limited (both spatially and semantically) by the quality of the metaanalytic data in the automaticallyextracted Neurosynth database
- Output of the GCLDA model is necessarily data and contextdependent, and model is based on a simplifying onetoone mapping between semantic representations and brain regions
- Interpreting scores: lack of consistent reporting standard for decoding between researchers

Summary

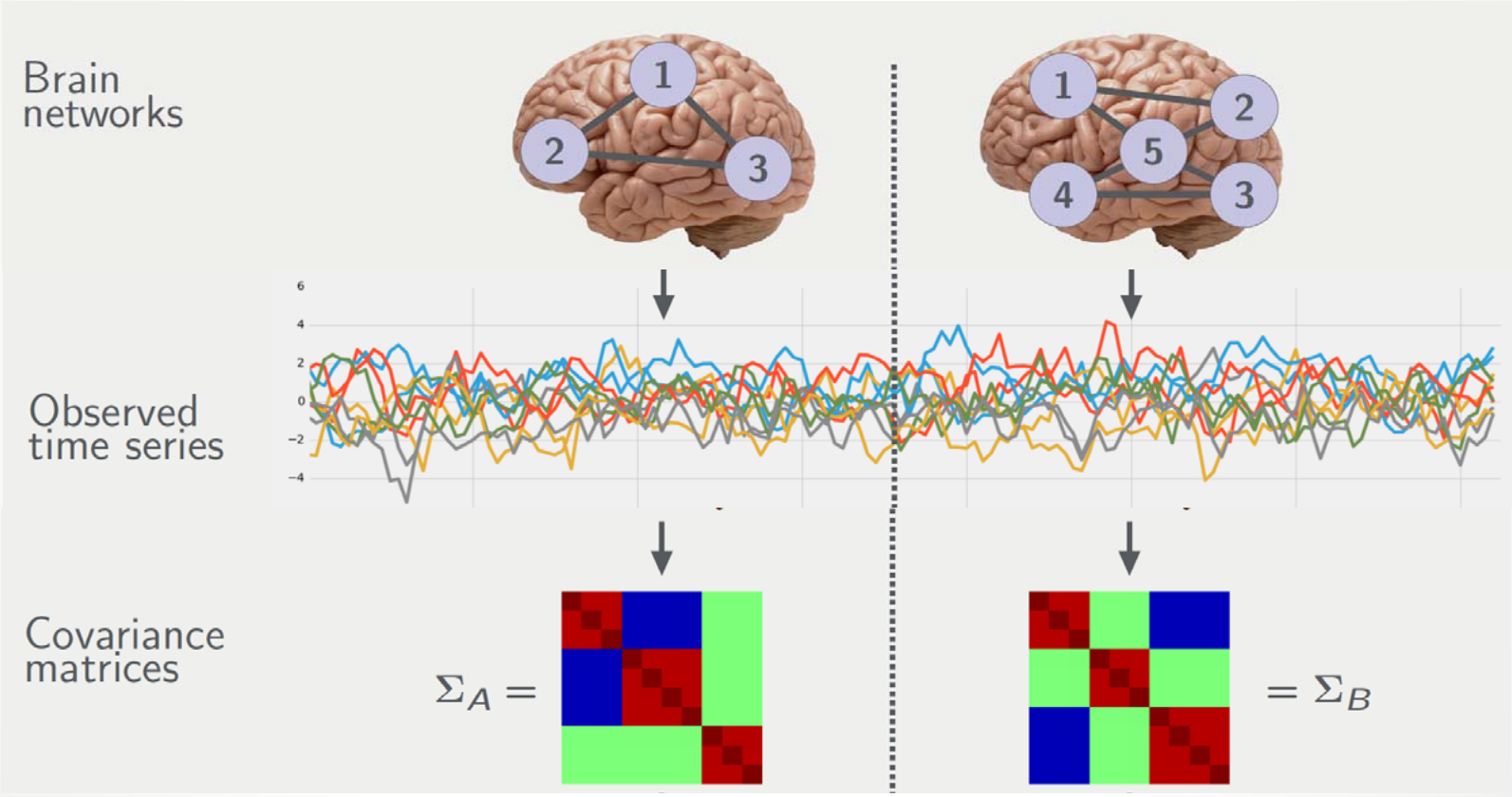
- Open-ended encoding & decoding are important for understanding the relationship between brain function and cognitive processes
- Proposed the GC-LDA model for learning a joint latent representation
- Extensive empirical validation of proposed model, applications include:
  - encoding & decoding (both fine grained and broad)
  - exploratory analysis
  - contextual predictions
  - summary & reconstruction
- Estimated topics maps are available in Neurovault

Part 2:
A Hierarchical Model for Time-Varying Functional Connectivity

Michael Riis Andersen @ Technical University of Denmark

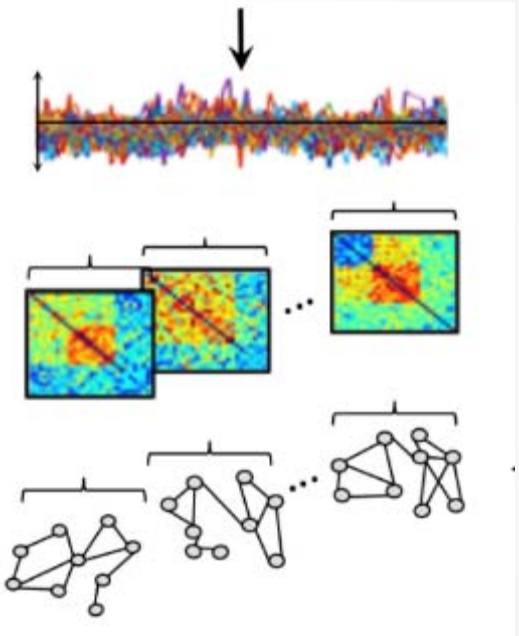
Russ Poldrack @ Stanford

Time Varying Connectivity



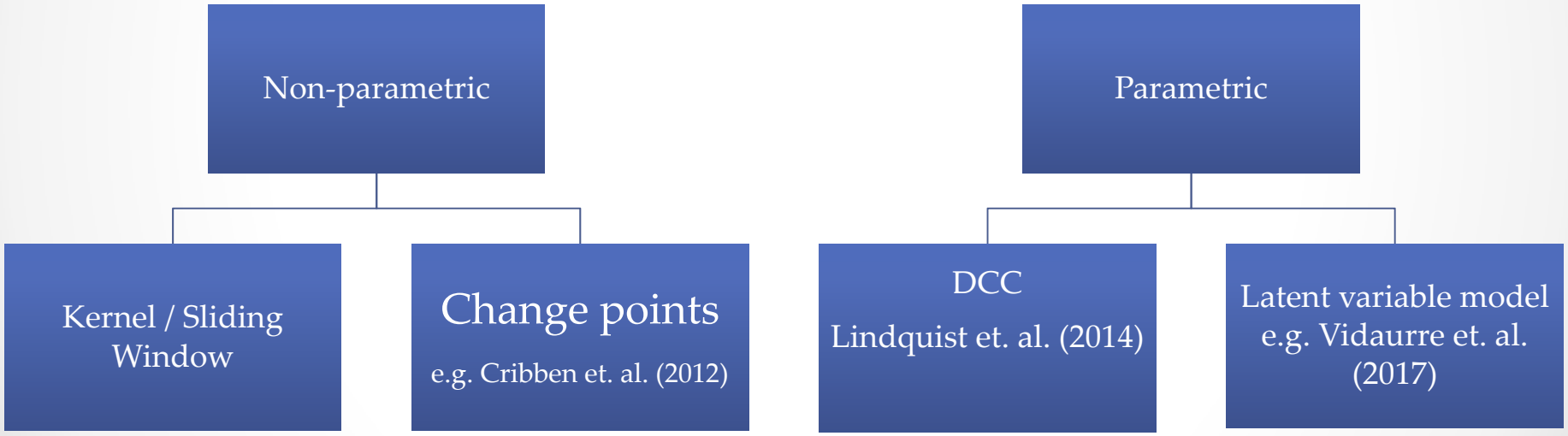
Motivating Questions

- How are the regions of the brain functionally connected?
- How do these connections change over time?
- ...
- How are the changing connections related to behavior, disease, etc.?

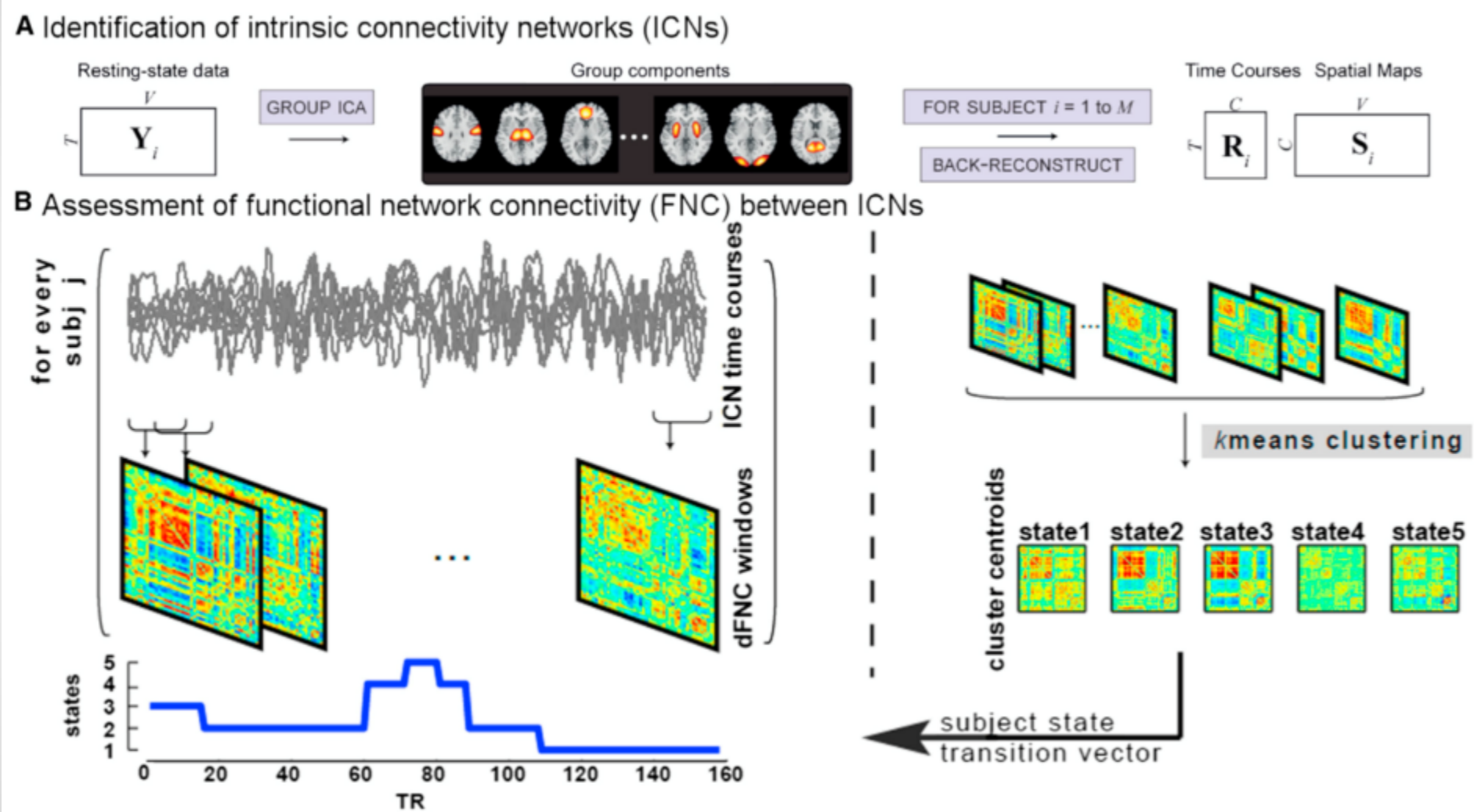


Shine et. al. (2016)

Estimating Time-Varying Functional Connectivity



Sliding Windows



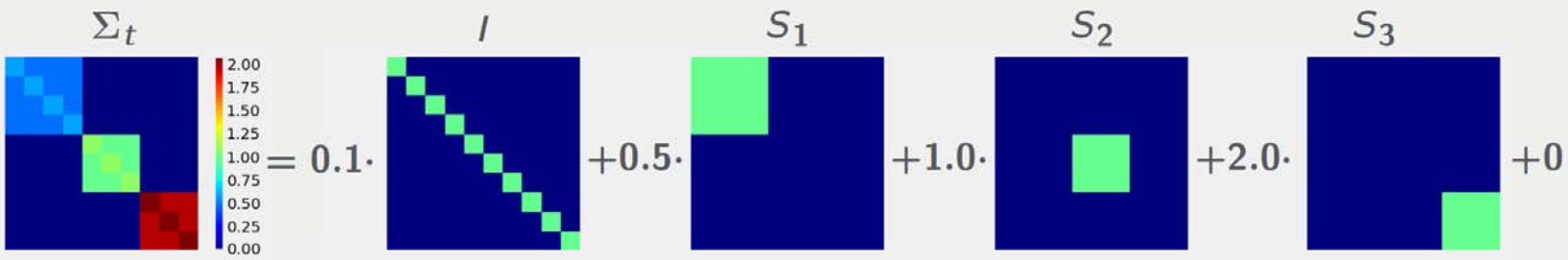
Calhoun et. al. (2015)

Open Issues

- Latent states may not be discrete
- Window size significantly affects results
- Separate window estimation can only use information within window

Model Hypotheses

- Subnetworks combined dynamically over time



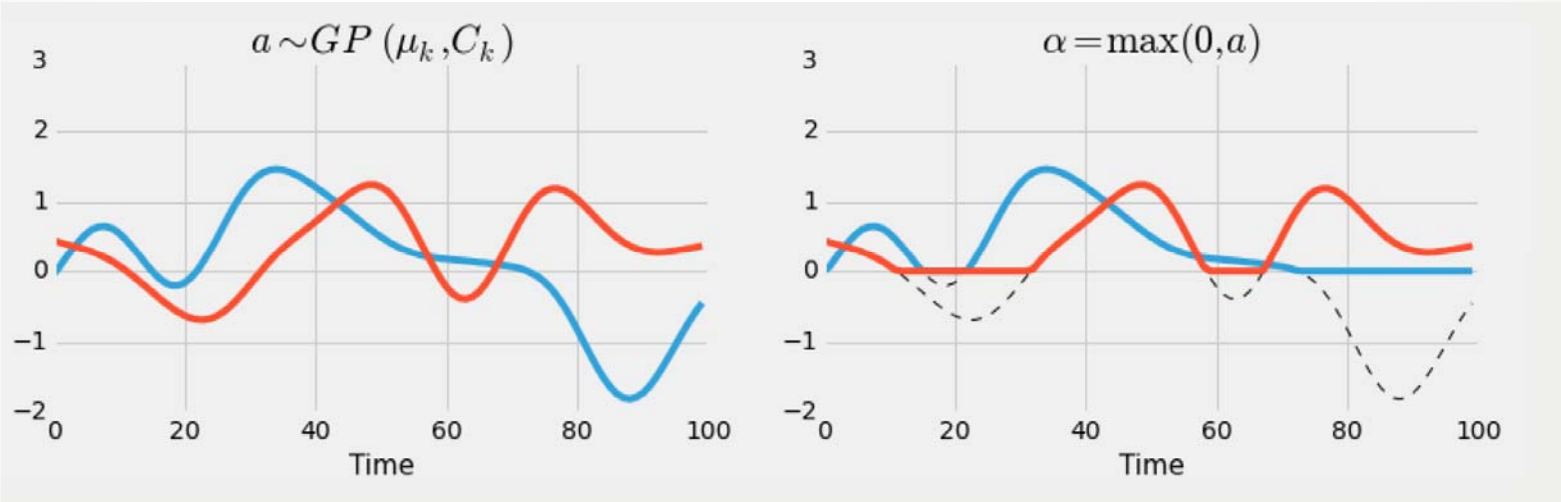
$$\Sigma_t = \beta^{-1} I + \sum_{k=1}^{K} \alpha_{k,t} S_k,$$

- Subnetworks should be "simple"

$$S_k = v_k v_k^T,$$
$$p(v_{k,i}) = (1 - p_k)\delta(v_{k,i}) + p_k \mathcal{N}\left(v_{k,i} \middle| 0, \tau_k\right)$$

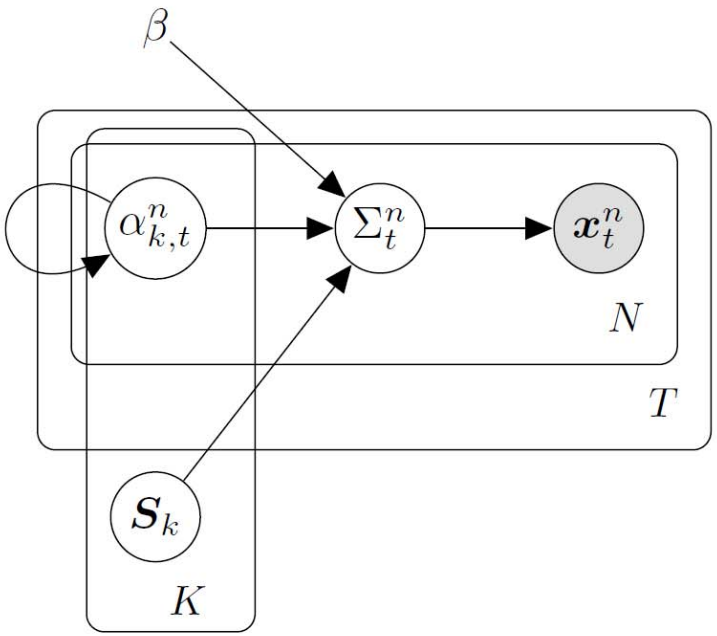
Model Hypotheses

- Dynamics are continuous, sparse, and smoothly varying over time

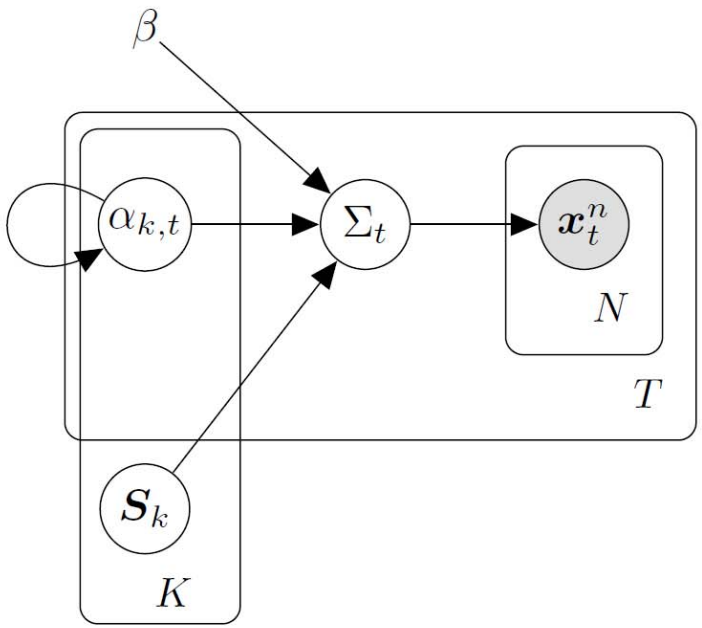


$$a_{k,\cdot} \sim \mathcal{GP}(\mu_k, C_k), \quad \alpha_{k,t} = \max(0, a_{k,t})$$

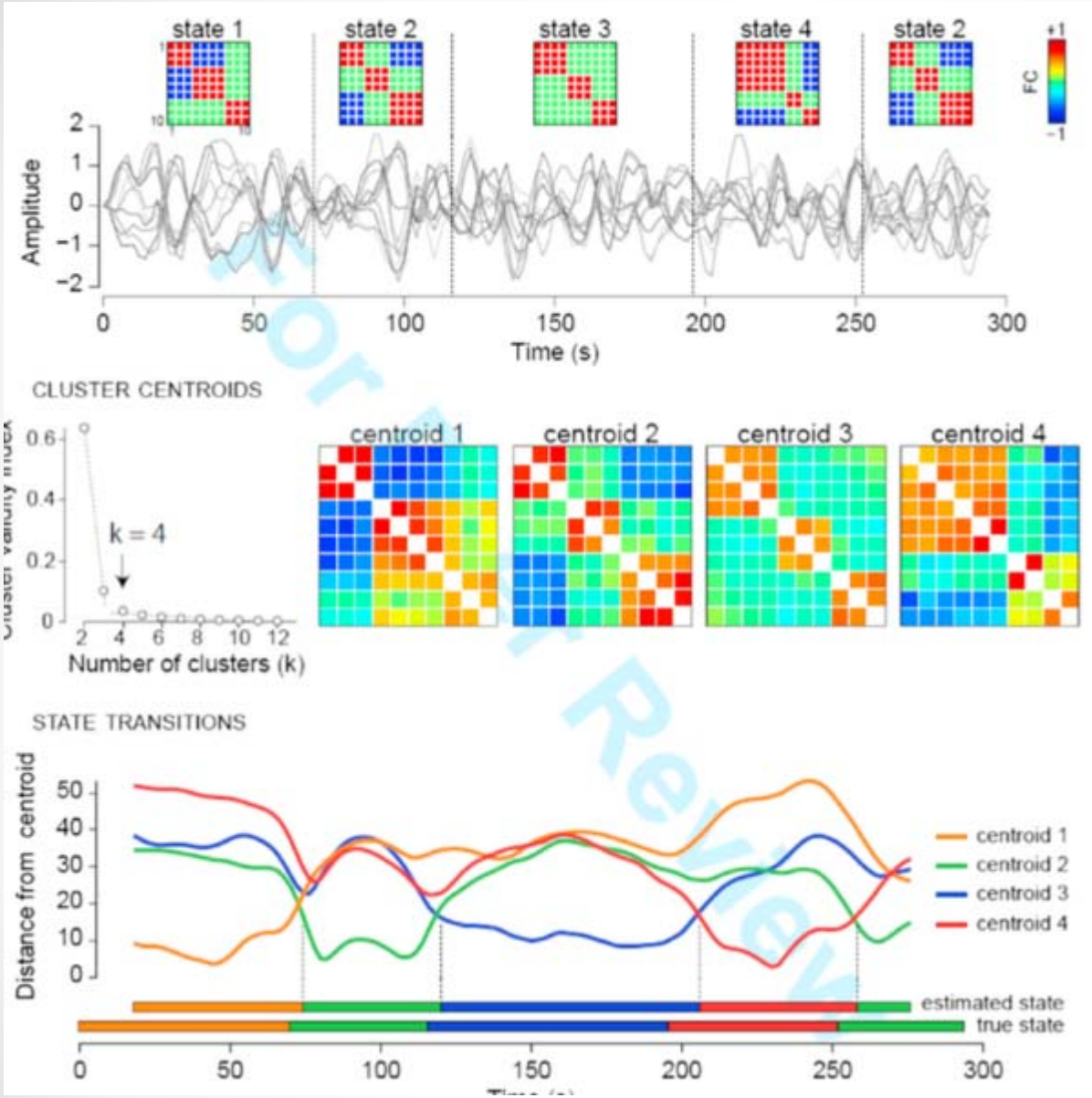
Proposed Models



(a) Subject specific mixing weights

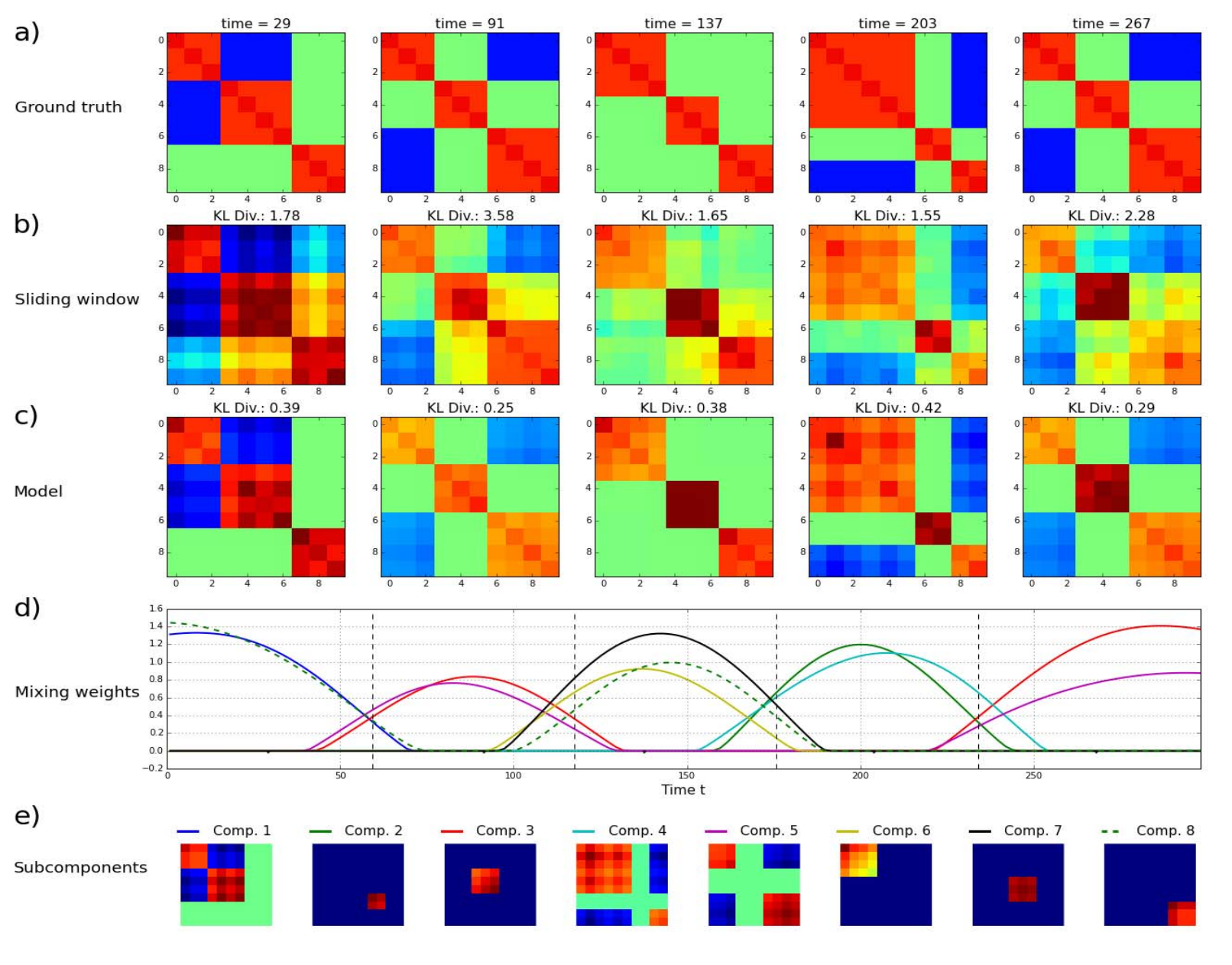


(b) Shared mixing weights across subjects

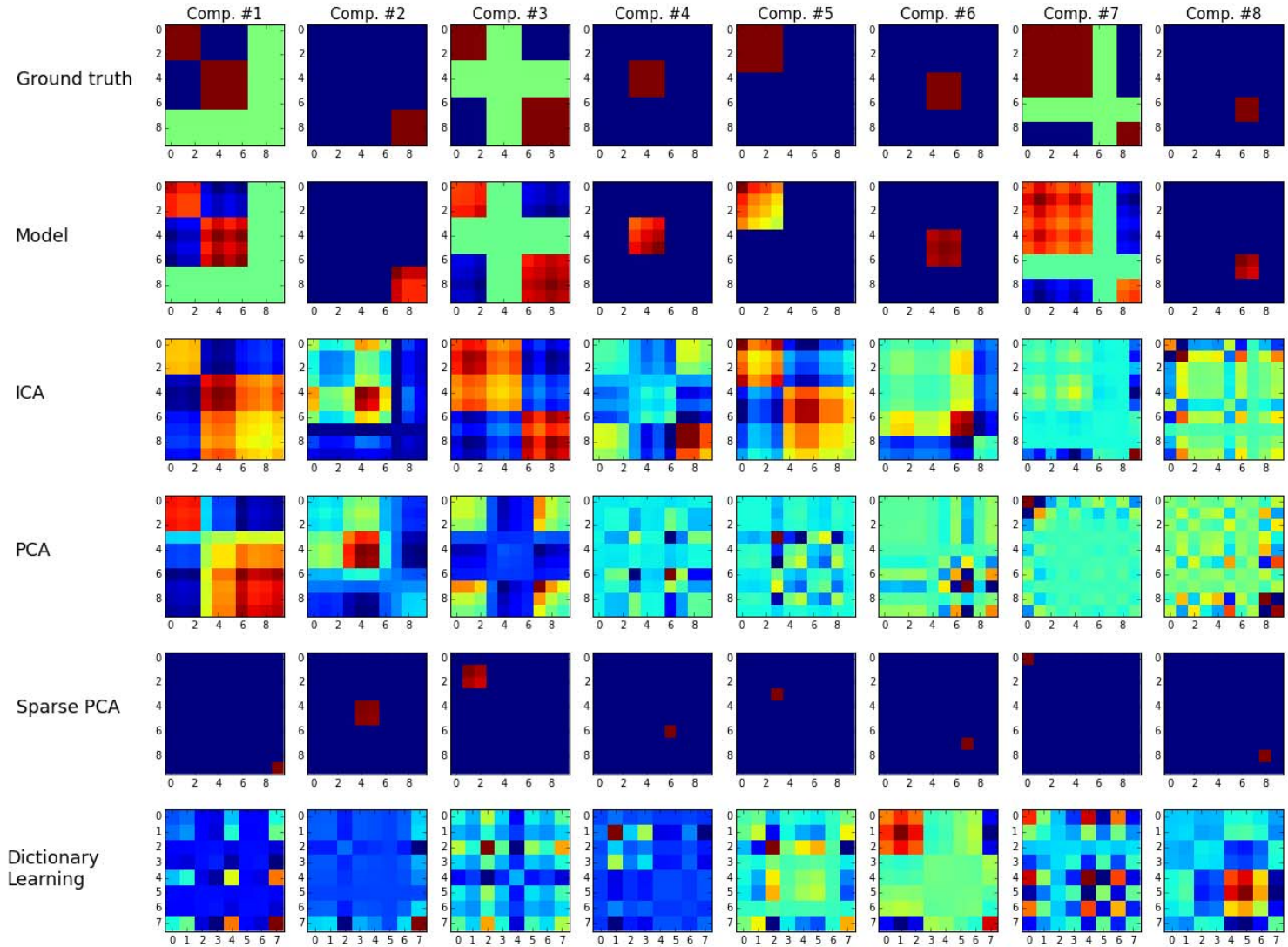


Simulated time-varying Connectivity (SimTB toolbox)

Calhoun et. al. (2015)

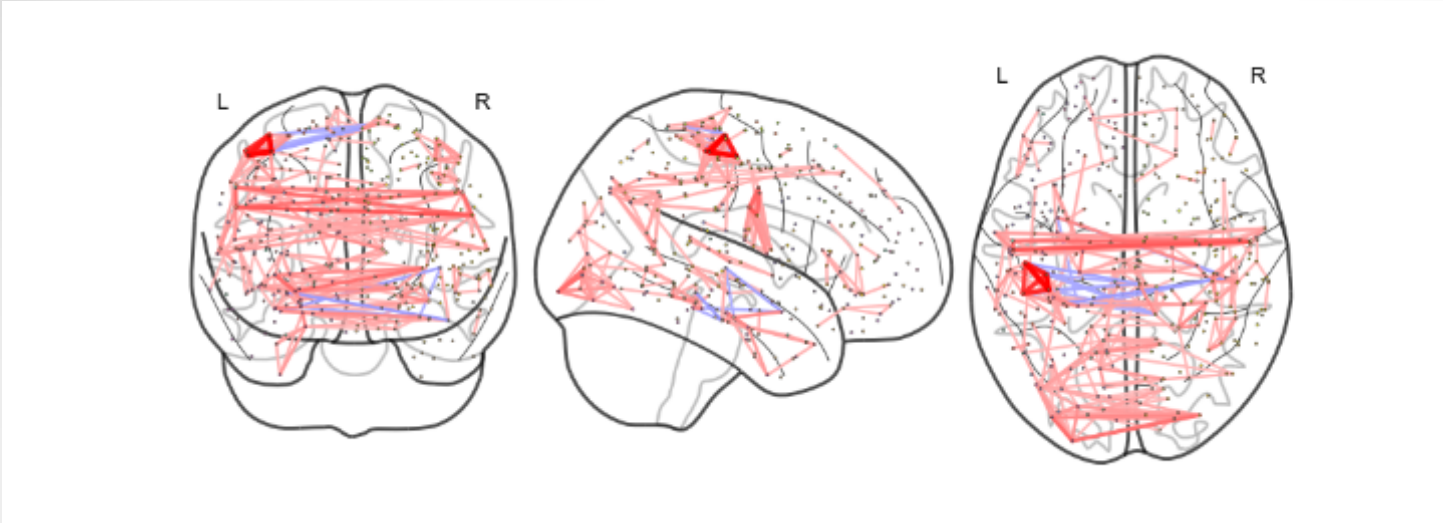

a)
Ground truth
time = 29
time = 91
time = 137
time = 203
time = 267
b)
Sliding window
KL Div.: 1.78
KL Div.: 3.58
KL Div.: 1.65
KL Div.: 1.55
KL Div.: 2.28
c)
Model
KL Div.: 0.39
KL Div.: 0.25
KL Div.: 0.38
KL Div.: 0.42
KL Div.: 0.29
d)
Mixing weights
Time t
e)
Subcomponents
Comp. 1
Comp. 2
Comp. 3
Comp. 4
Comp. 5
Comp. 6
Comp. 7
Comp. 8

Factor Model Visualization

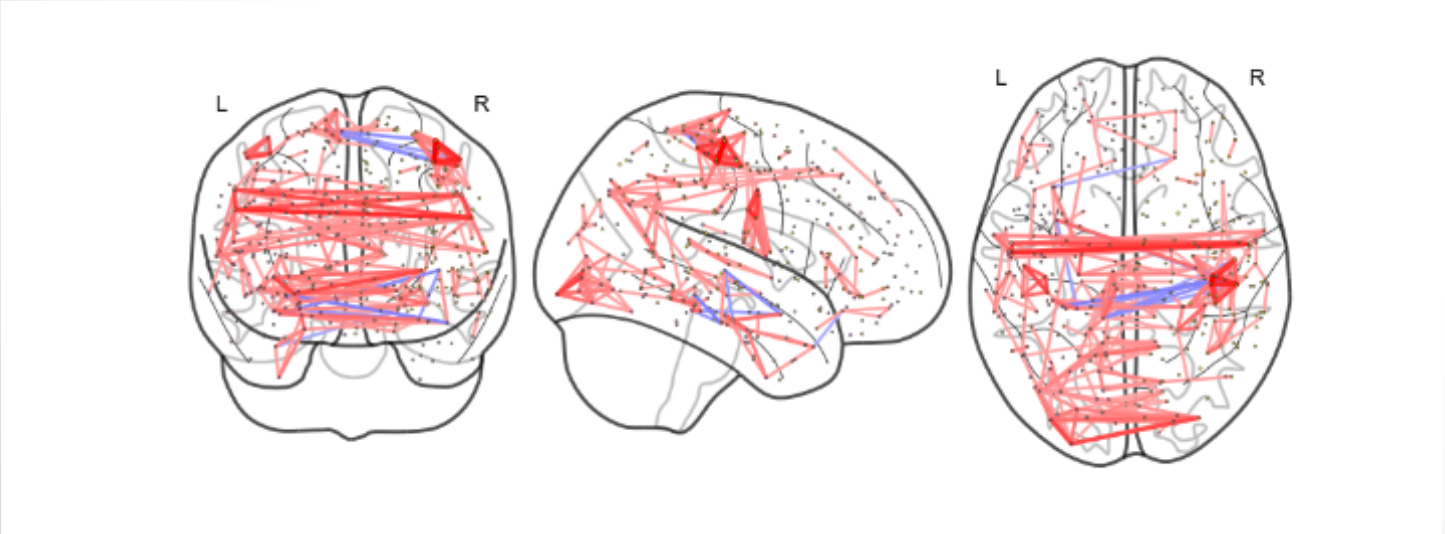


Evaluation

- HCP data, Gordon 333 atlas, Motor task
  - Right/left hand tapping, right/left foot tapping, tongue wagging, rest
- Task block + motion regressed out, model the residual
- Evaluate using cross-validation for predicting task on held-out subjects
- Compared to sliding window

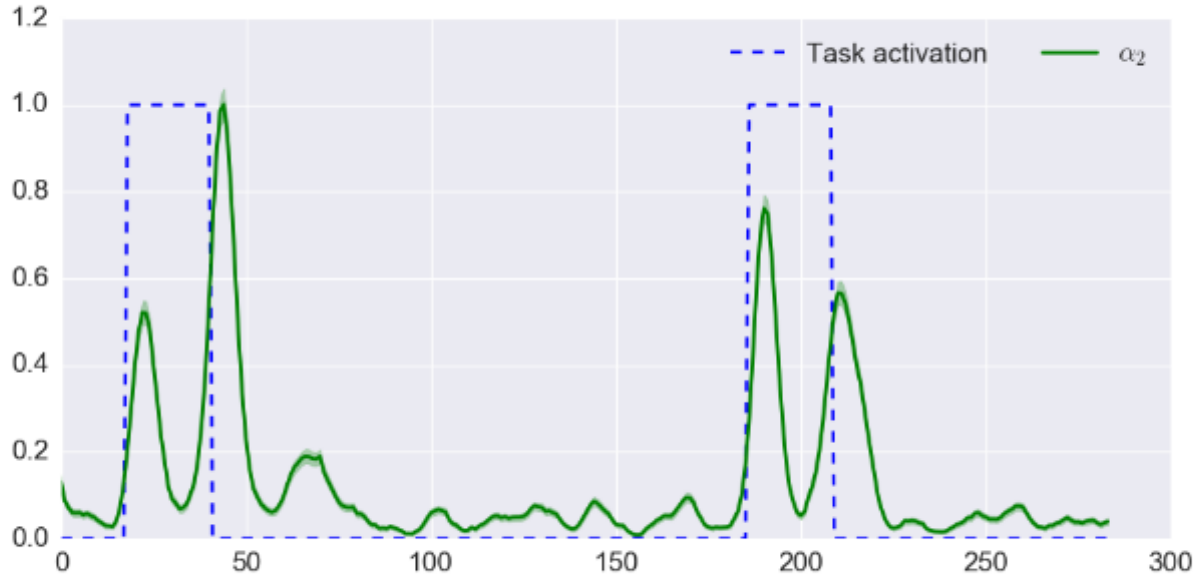


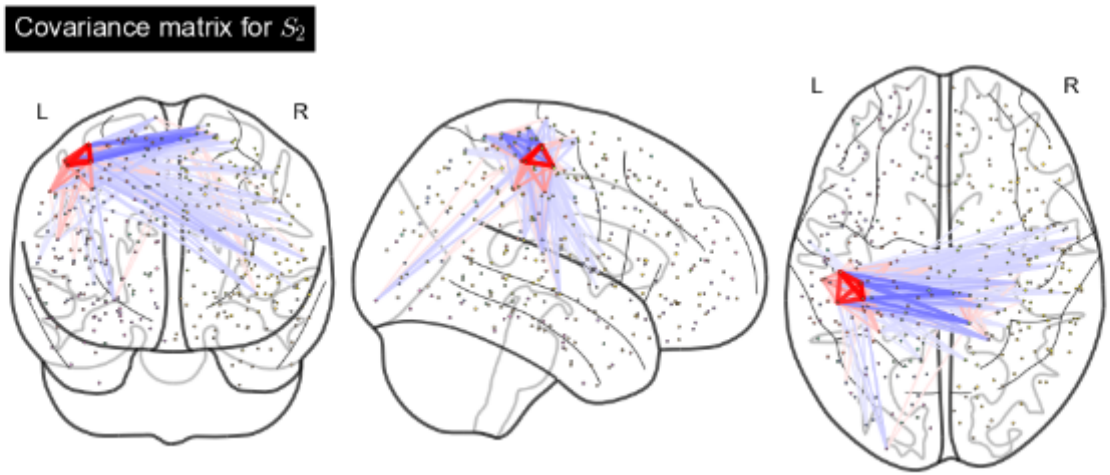
Right Hand Tapping

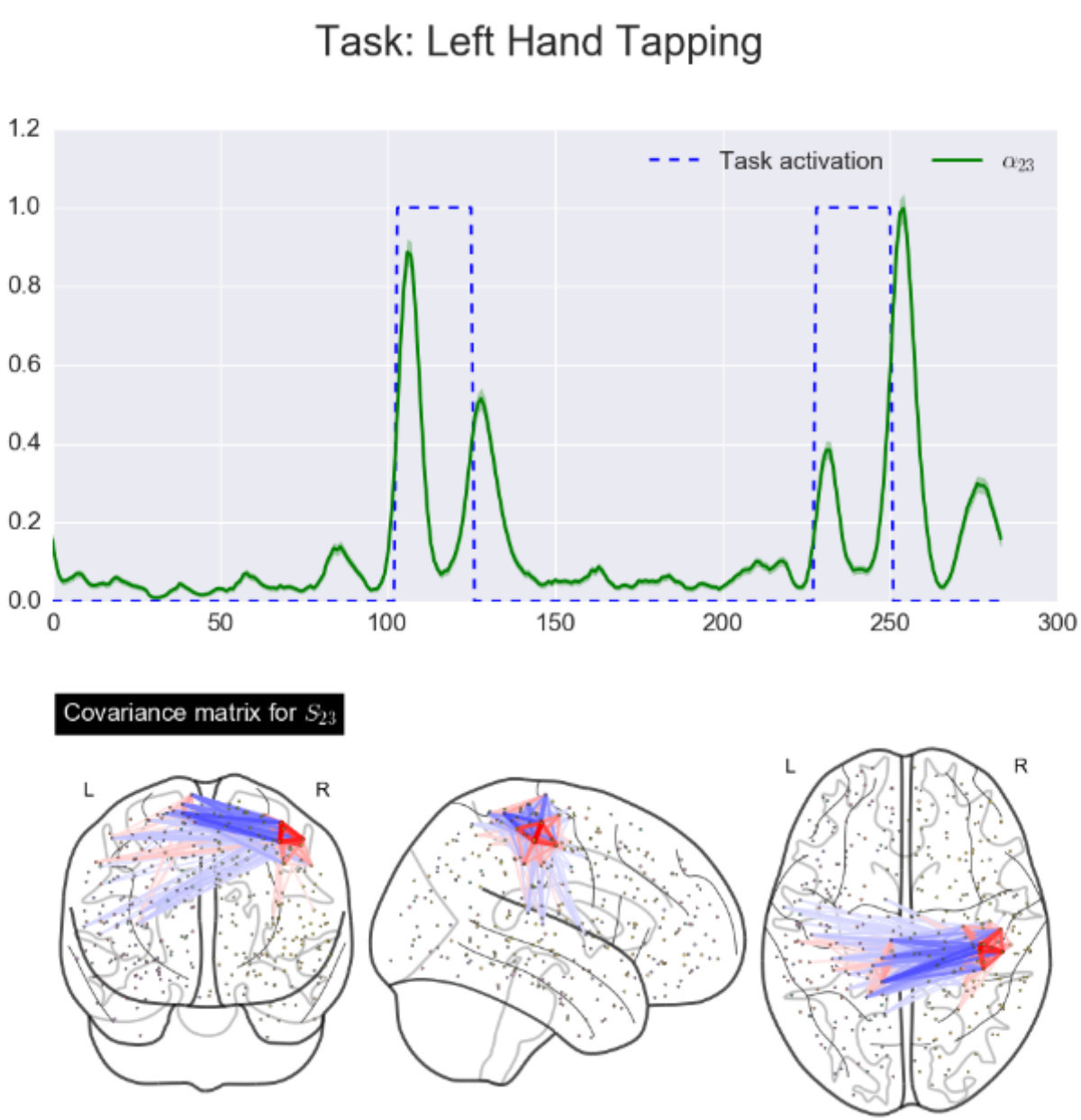


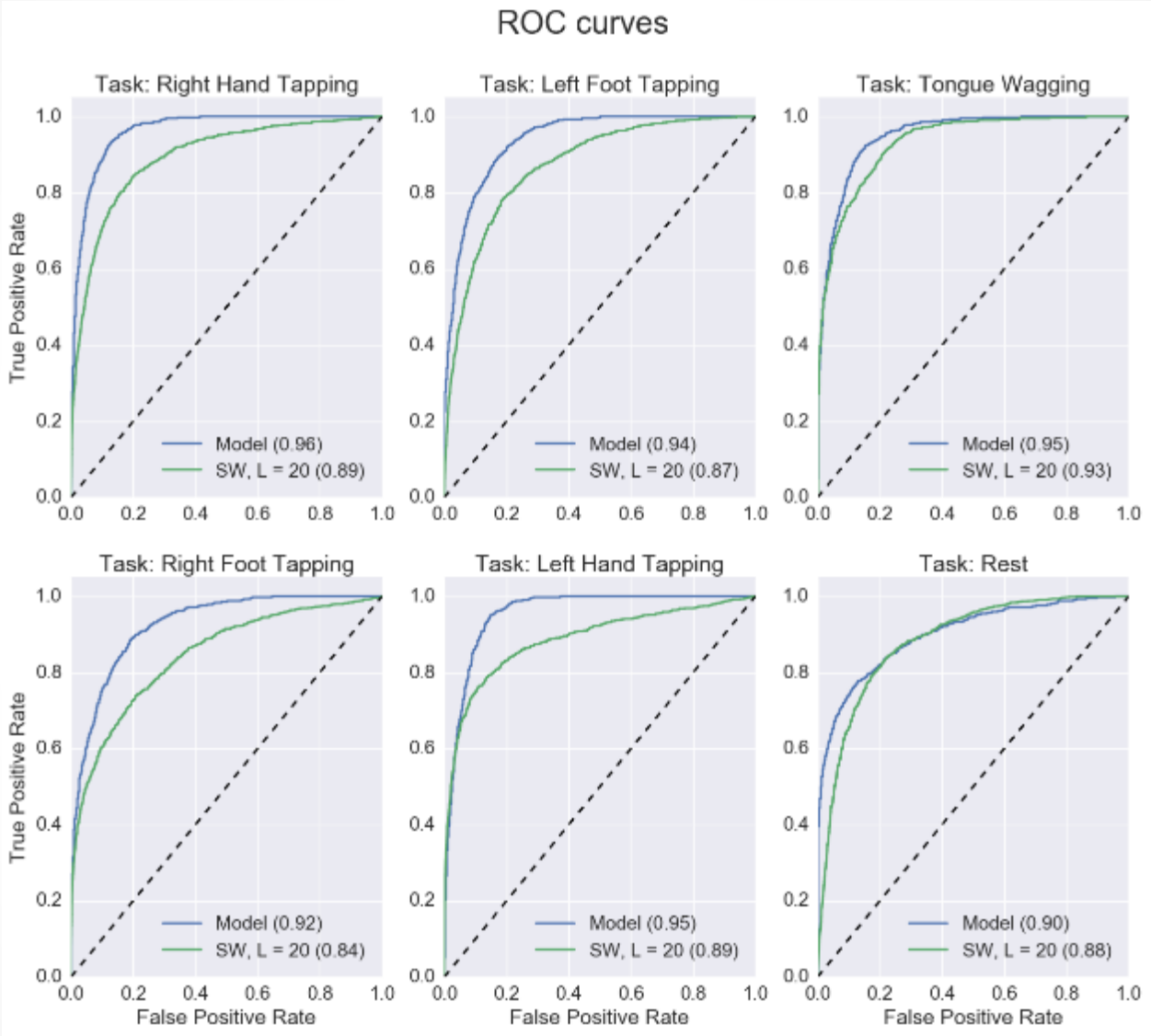
Left Hand Tapping

Task: Right Hand Tapping






Task: Left Hand Tapping
Task activation
α_{23}
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
250
300
Covariance matrix for S_{23}
L
R
L
R


ROC curves
Task: Right Hand Tapping
Model (0.96)
SW, L = 20 (0.89)
Task: Left Foot Tapping
Model (0.94)
SW, L = 20 (0.87)
Task: Tongue Wagging
Model (0.95)
SW, L = 20 (0.93)
Task: Right Foot Tapping
Model (0.92)
SW, L = 20 (0.84)
Task: Left Hand Tapping
Model (0.95)
SW, L = 20 (0.89)
Task: Rest
Model (0.90)
SW, L = 20 (0.88)
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0

Summary (Part 2)

- Improved model for time varying connectivity that captures scientific intuition, e.g. spatial sparsity and temporal smoothness
- Empirical validation of proposed model on HCP, better classification of task states
- Work in Progress:
  - Improving computational scalability
  - Predictive evaluation for resting state data
  - Incorporating spatial structure (particularly for high resolution)

Overall Summary

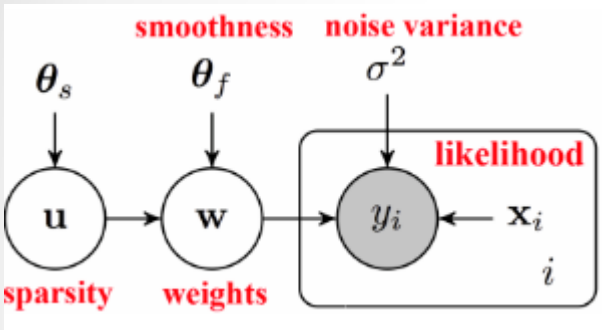
- Model-based approach provides a straightforward mechanism for synthesizing, evaluating and comparing scientific hypotheses
- Shown applications to decoding and time varying connectivity
- Multiple open-source tools for (almost) math-free inference e.g. STAN, infer.NET, …

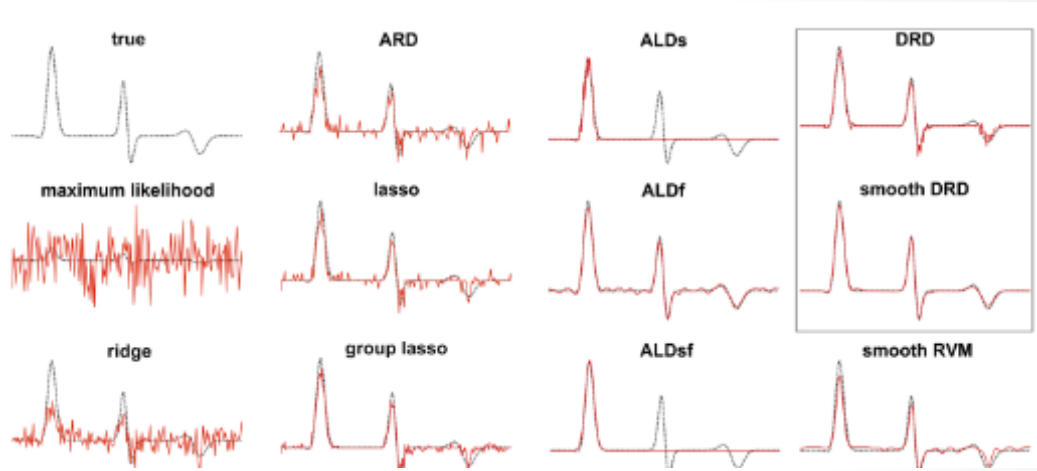
Research interests: developing methods for understanding the brain

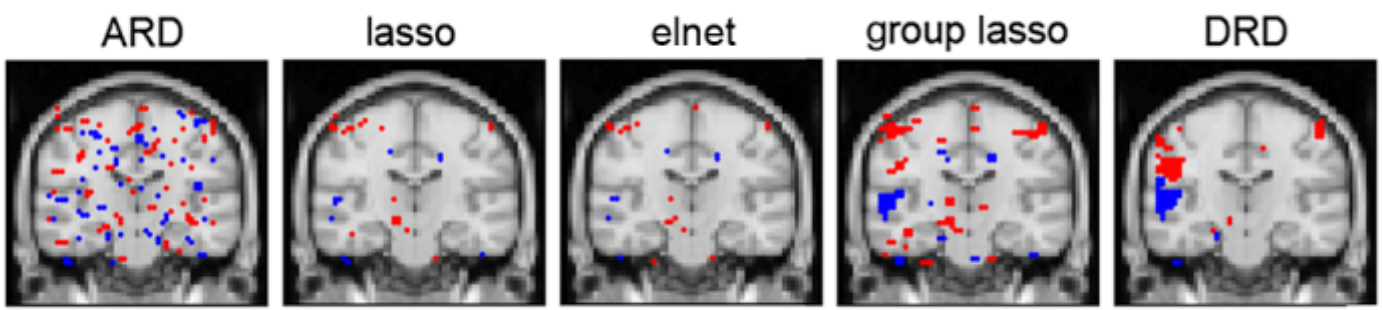
- Closed-world Interpretable Encoding and Decoding
- Open-Ended Encoding & Decoding (This Talk)
- Improved estimators for time varying connectivity (This Talk)
- Investigating relationships between connectivity and behavior / disease

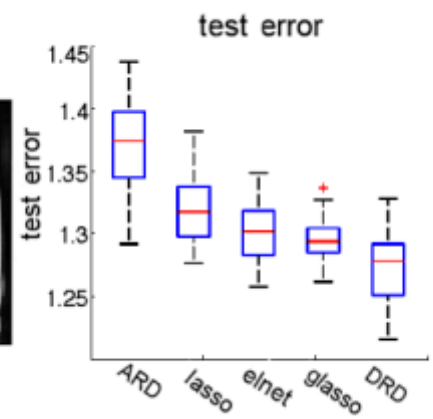
- Also, Research on Probabilistic & Statistical Machine Learning

Interpretable Models for Decoding



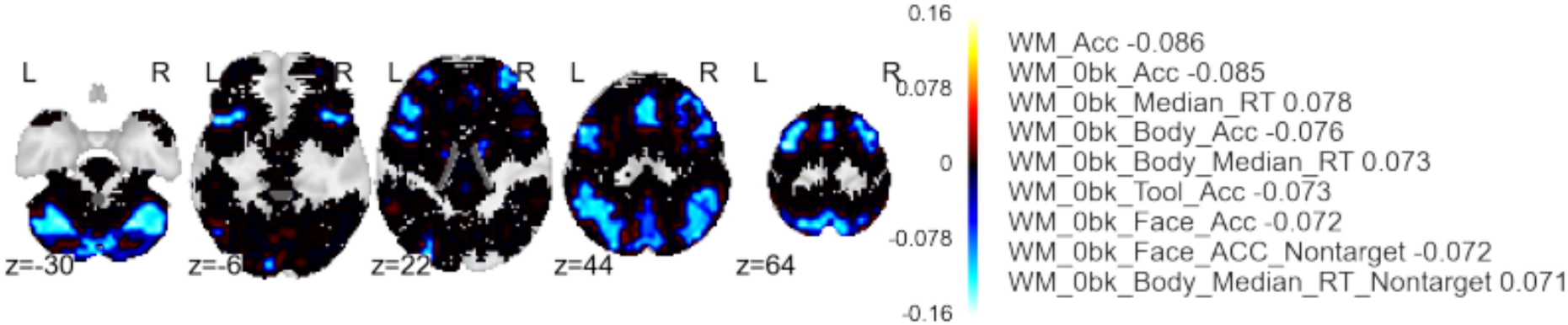


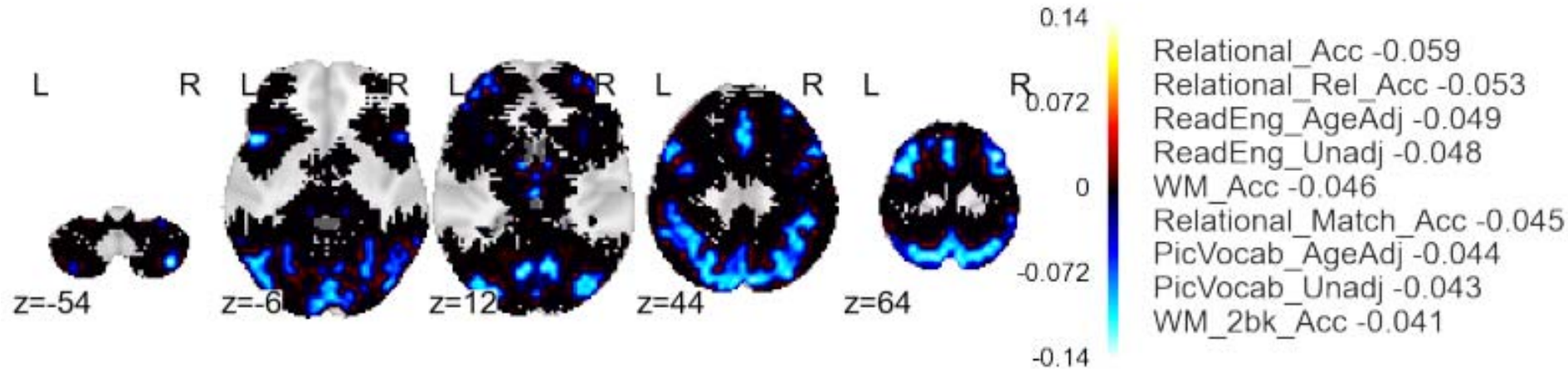




- HCP relational reasoning -> Penn Matrix task (Wu et. al, 2015)

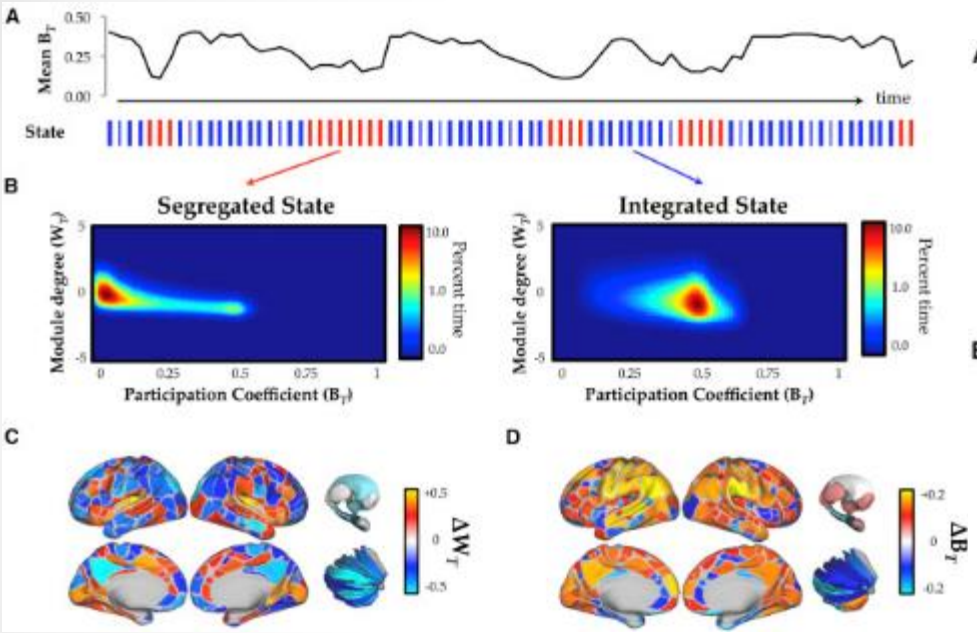
Sparse CCA for Imaging & Behavior

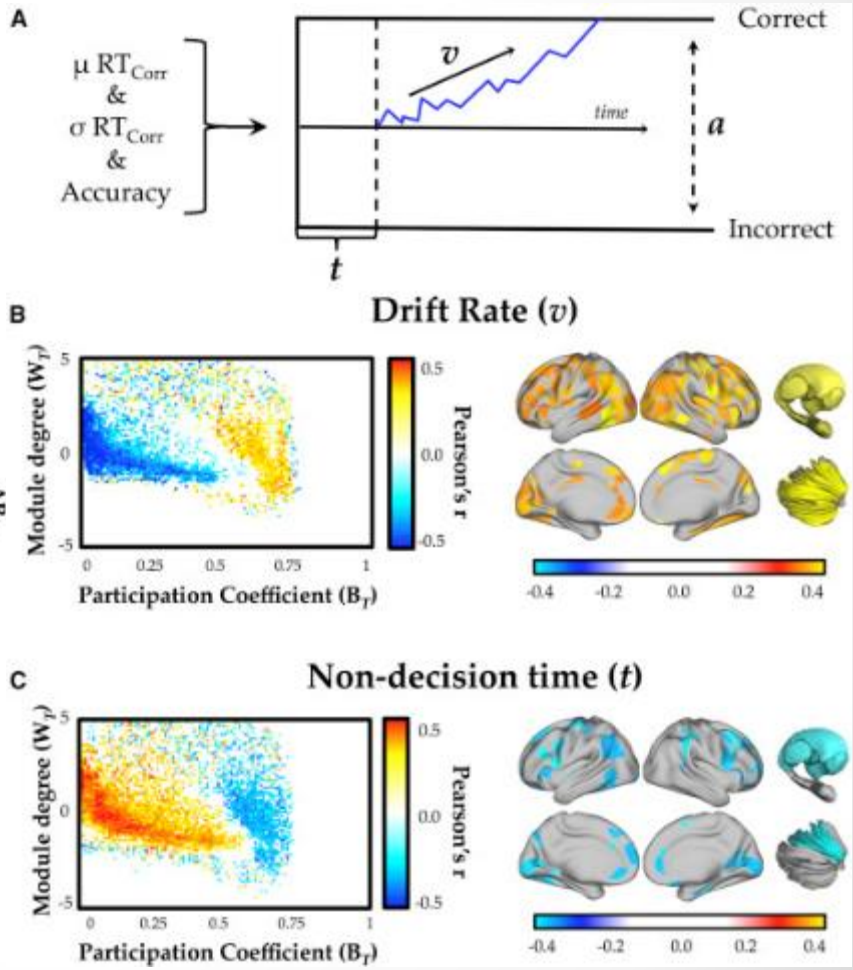




- HCP 2-back and Relational Reasoning (Khanna et. al. 2016)

Relationship between Connectivity and Behavior





Shine et. al. (2016)

Thank You!!!

Questions?

contact: sanmi@illinois.edu

References-I

- Rubin, T. N., Koyejo, O., Gorgolewski, K. J., Jones, M. N., Poldrack, R. A., & Yarkoni, T. (2016). Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition. bioRxiv, 059618
- Timothy Rubin, Oluwasanmi Koyejo, Michael Jones, and Tal Yarkoni (2016). Generalized Correspondence-LDA Models (GC-LDA) for Identifying Functional Regions in the Brain. In Advances in Neural Information Processing Systems (NIPS) 29, 2016
- Poldrack, R. A., Halchenko, Y. O., & Hanson, S. J. (2009). Decoding the large-scale structure of brain function by classifying mental states across individuals. Psychological Science, 20(11), 1364-1372.
- Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., & Greicius, M. D. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cerebral cortex, 22(1), 158-165.
- Gorgolewski, K. J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S. S., Maumet, C., ... & Yarkoni, T. (2015). NeuroVault. org: A web-based repository for collecting and sharing unthresholded statistical maps of the human brain. Frontiers in neuroinformatics, 9, 8.

References-II

- Fox, P. T., Mikiten, S., Davis, G., & Lancaster, J. L. (1994). BrainMap: a database of human functional brain mapping. Functional Neuroimaging: Technical Foundations, 95-105.
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. Nature methods, 8(8), 665-670.
- Chang, L. J., Yarkoni, T., Khaw, M. W., & Sanfey, A. G. (2012). Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. Cerebral Cortex, bhs065.
- Poldrack, R. A., Mumford, J. A., Schonberg, T., Kalar, D., Barman, B., & Yarkoni, T. (2012). Discovering relations between mind, brain, and mental disorders using topic mapping. PLoS Comput Biol, 8(10), e1002707.
- Helfinstein, S. M., Schonberg, T., Congdon, E., Karlsgodt, K. H., Mumford, J. A., Sabb, F. W., ... & Poldrack, R. A. (2014). Predicting risky choices from brain activity patterns. Proceedings of the National Academy of Sciences, 111(7), 2470-2475.

References-III

- Yeo, B. T., Krienen, F. M., Eickhoff, S. B., Yaakub, S. N., Fox, P. T., Buckner, R. L., ... & Chee, M. W. (2014). Functional specialization and flexibility in human association cortex. Cerebral Cortex.
- Laird, A. R., Eickhoff, S. B., Fox, P. M., Uecker, A. M., Ray, K. L., Saenz, J. J., ... & Turner, J. A. (2011). The BrainMap strategy for standardization, sharing, and meta-analysis of neuroimaging data. BMC research notes, 4(1), 349.

References – Prior Work

- Wu, A., Park, M., Koyejo, O. O., & Pillow, J. W. (2014). Sparse Bayesian structure learning with "dependent relevance determination" priors. In Advances in Neural Information Processing Systems (pp. 1628-1636).
- Khanna, R., Ghosh, J., Poldrack, R., & Koyejo, O. (2016). Information Projection and Approximate Inference for Structured Sparse Variables. arXiv preprint arXiv:1607.03204.
- Andersen, J. M., Koyejo, O., & Poldrack, R. A. (2016). Model-based dynamic resting state functional connectivity. Under preparation, OHBM 2016 poster
- Shine, J. M., Bissett, P. G., Bell, P. T., Koyejo, O., Balsters, J. H., Gorgolewski, K. J., ... & Poldrack, R. A. (2016). The Dynamics of Functional Brain Networks: Integrated Network States during Cognitive Task Performance. Neuron.

Replication of Poldrack et al. (2012)

A Topic 71 (115 docs):
narrative, discourse, comprehension, memory, discourse_processing

Topic 61 (441 docs):
memory, working_memory, maintenance, visual_working_memory, spatial_working_memory

B Topic 33 (479 docs):
sentences, language, comprehension, sentence, processing

Topic 22 (485 docs):
memory, working, verbal, term, load, maintenance, performance