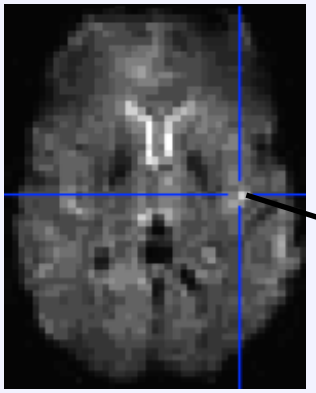
Large-scale automated synthesis of human functional neuroimaging data.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). *Nature Methods*, 8(8), 665–670.

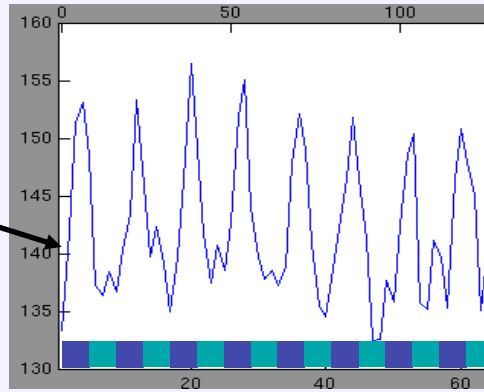
Presentation by Thomas Nichols, PhD
Department of Statistics, Warwick Manufacturing Group
University of Warwick

Understanding the Brain with Functional Neuroimaging

- ~20 years of “functional neuroimaging”
  - Functional Magnetic Resonance Imaging (fMRI)
  - Measures changes in blood flow
- Cognitive Neuroscience
  - Correlating brain activity with cognition, behavior, emotion, pain, etc.
  - “Brain Mapping”



Single fMRI Image

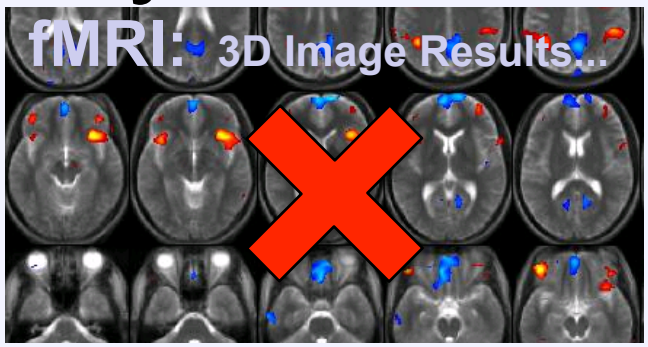


fMRI Time Series

Meta-Analysis for Brain Imaging

- Meta-Analysis
  - Integration of results from multiple studies
    - To increase power from pooling studies
    - To discover which results generalize, which are idiosyncratic
- fMRI – an explosion of studies
  - 2 papers in 1993 → 2446 papers in 2010
  - Sample sizes small (10-20 subjects)
- Great need for fMRI meta-analysis
  - But several challenges…

Imaging Meta-Analysis



- Authors don't share their image data
- Only coordinates available
  - x,y,z peaks of activation maps
    - In standard brain atlas coordinates
- Meta-analysis
  - x,y,z peak manually collected from papers

x, y, z atlas coordinates

Activations and deactivations with maternal love

	Left				Right			
	x	*y*	*z*	*Z*	*x*	*y*	*z*	*Z*
Activations (Brodmann area)								
Middle insula (14)*	−42	8	−4	3.97	46	12	−6	3.22
Anterior cingulate dorsal (24)*	−10	26	30	2.85	10	34	18	2.75
Anterior cingulate ventral (24)	−8	44	−6	3.89	10	44	−6	1.87
Caudate nucleus	−22	2	16	2.37	14	−2	18	2.67

Results Table of Coordinates

Standard reporting format for a fMRI publication

Imaging Meta-Analysis: Problem

- Manual 'harvesting of points'
  - x,y,z peak **manually** collected from papers
  - Topic of paper **manually** scored
    - Vision study? Emotion study? Memory study?
  - Prone to error, limits number of studies considered

x, y, z atlas coordinates

Activations and deactivations with maternal love

	Left				Right			
	x	*y*	*z*	*Z*	*x*	*y*	*z*	*Z*
Activations (Brodmann area)								
Middle insula (14)*	−42	8	−4	3.97	46	12	−6	3.22
Anterior cingulate dorsal (24)*	−10	26	30	2.85	10	34	18	2.65
Anterior cingulate ventral (24)	−8	44	−6	3.89	10	44	−6	1.87
Caudate nucleus	−22	2	16	2.37	14	−2	18	2.67

Results Table of Coordinates

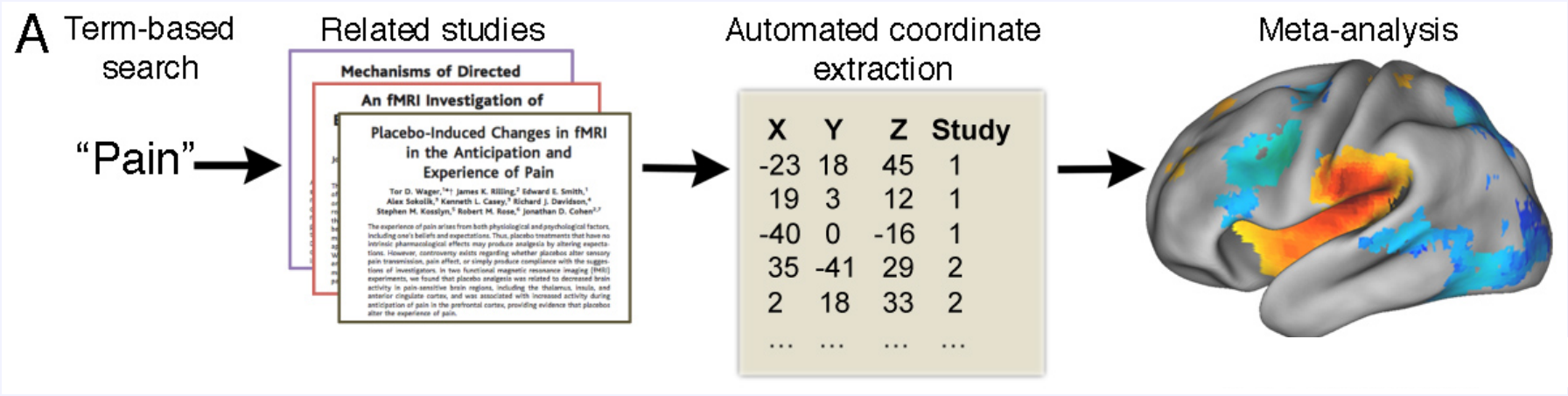
Standard reporting format for a fMRI publication

“Neurosynth” – automated neuroimaging meta-anlaysis

- Uses HTML versions of published papers
- From each paper…
  - Extracts x,y,z coordinates
  - Records word frequency
    - Occurrence per 1000 words
      - e.g. “pain” occurs 6 times per 1000 words
- Then can relate locations to topics
  - e.g. does a given brain area associate more with “pain” or “memory”?

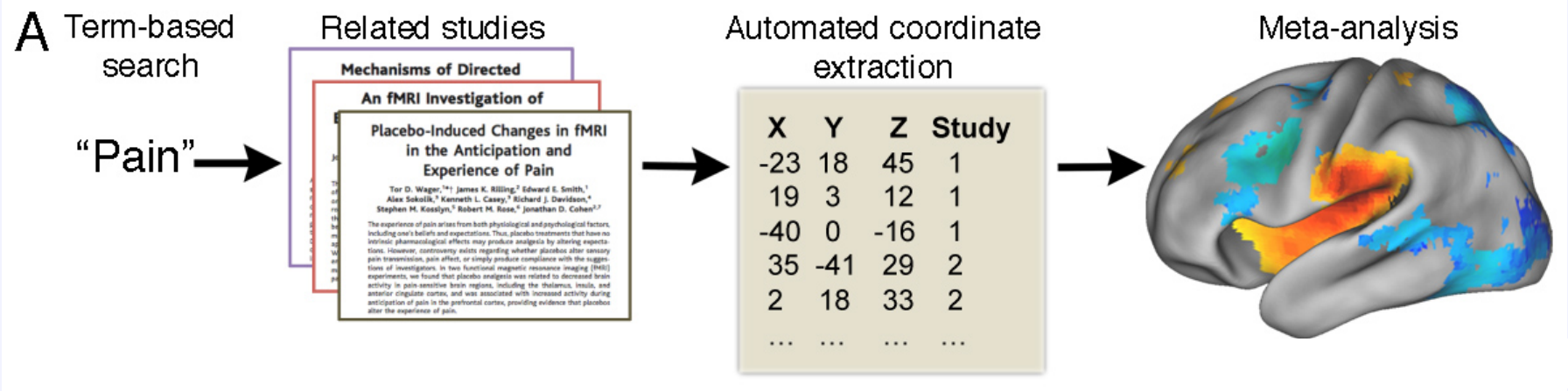
Key Details

- Data sample
  - 17 journals
  - 3,489 articles
  - 100,953 x,y,z coordinate locations
  - 10,000 words



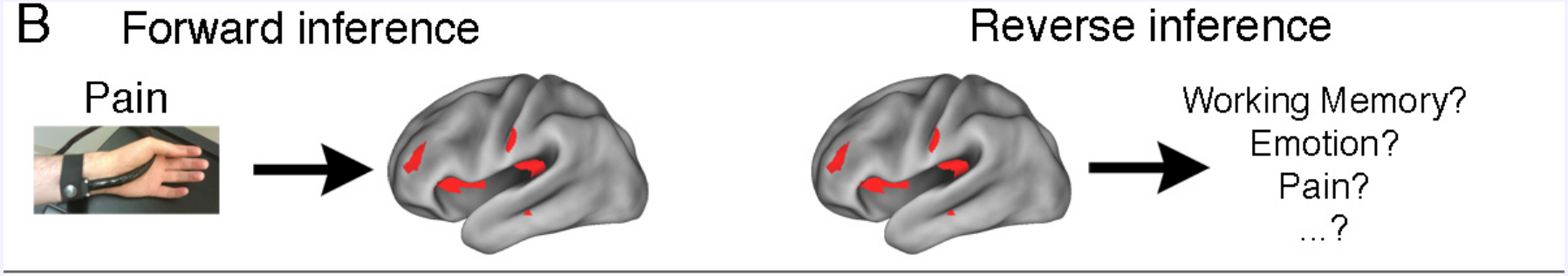
Key Details

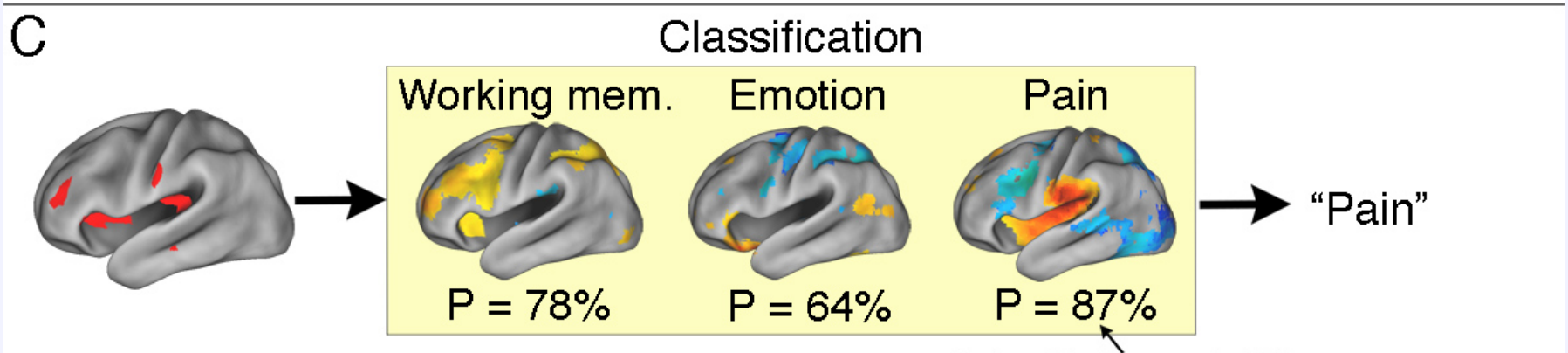
- For study i
  - Brain activation present (w/in 10 mm) or not
    $A_{ij} \in \{0,1\}$
    at voxel j
  - For term k present (> 1 in 1000 words) or not
    $T_{ik} \in \{0,1\}$
- $P(A|T_k=1)$ gives "forward inference"
  - Probability of activation in studies with term k



Key Details

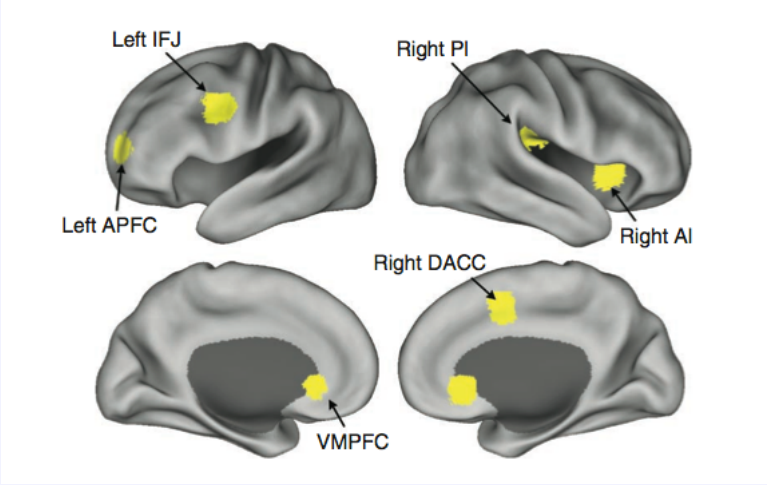
- P(T_k=1|A) gives "reverse inference"
  - Given a new study with activation A, predict term k associated with study
  - Naïve Bayes
    - Independence assumed between brain voxels
  - "Smoothing" used – add one "hit", one "miss"
    - Corresponds to Bernoulli MAP estimator





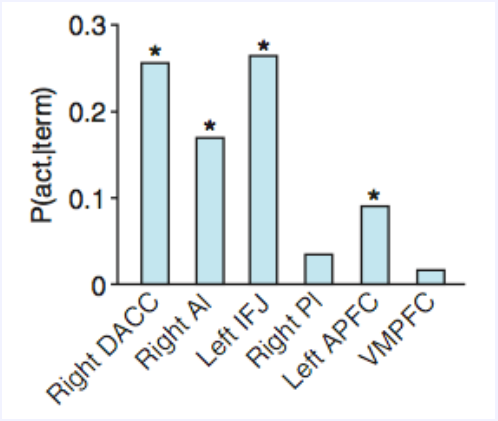
Key Results

- Forward & Reverse Inferences
  - Some areas weakly associated (forward), but highly specific (reverse)
  - Some areas strongly associated (forward), but not specific (reverse)

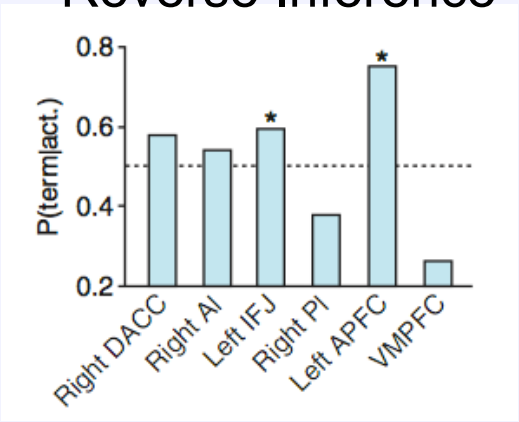


Term: "pain"

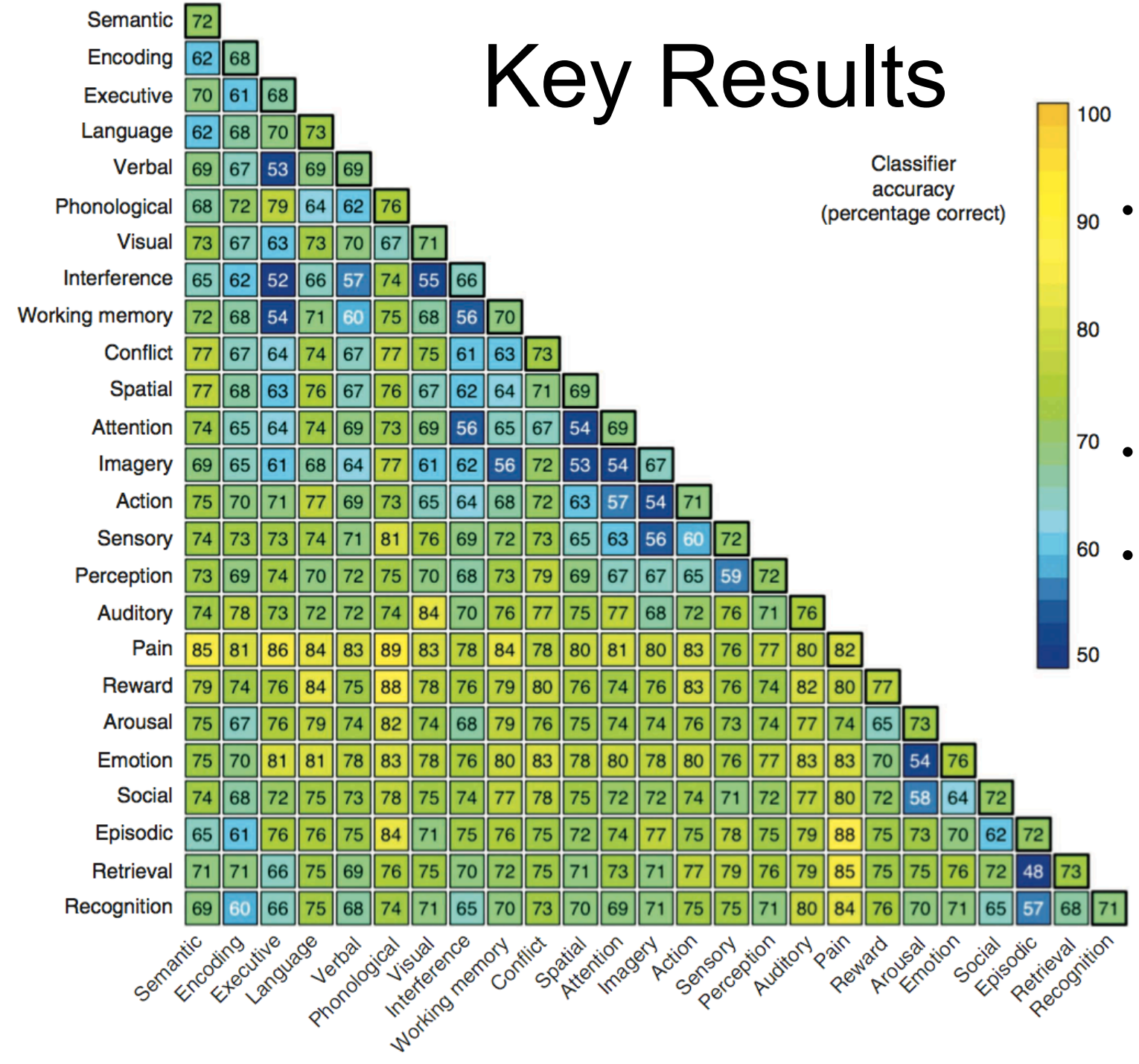
Forward Inference



Reverse Inference



Key Results



- Pairwise term classification
  - Based on coordinate data
- “pain” most different
- “Attention” & “Interference” very similar

Limitations

- Coordinate harvesting imperfect
  - Difficult to parse all HTML tables
- Terms are not true “topics”
  - An author can use many terms not directly about their own experiment
- Naïve Bayes
  - Voxels clearly dependent

Conclusions

- Automated neuroimaging meta-analysis
- Applies text-mining and classification techniques, to…
  - Speed and automate data collection
  - Provide “reverse” inferences
  - Explore relationship between different terms in literature
- On line tool
  - http://www.neurosynth.org