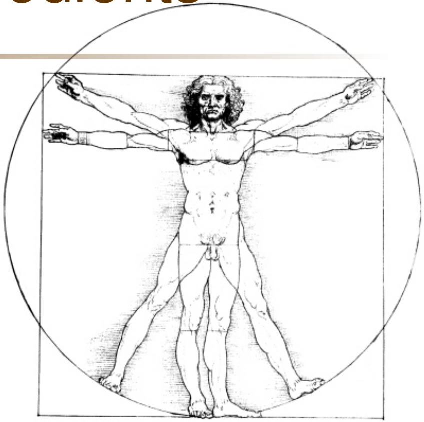
Model predictions advance science more than modeling ingredients

Gaël Varoquaux



AI as statistical methods for imperfect theories
[Varoquaux 2021]

My scientific wanderings

Physics

- Quantum physics (PhD with Alain Aspect)
  Atom-interferometric tests of relativity

Brain image analysis for cognition

- Statistics, machine learning, image analysis
- Cognitive neuroscience, psychology

Machine learning for public health

Informing policy?

From absolute quantities
to qualitative subject matters

How does scientific knowledge emerge from data?

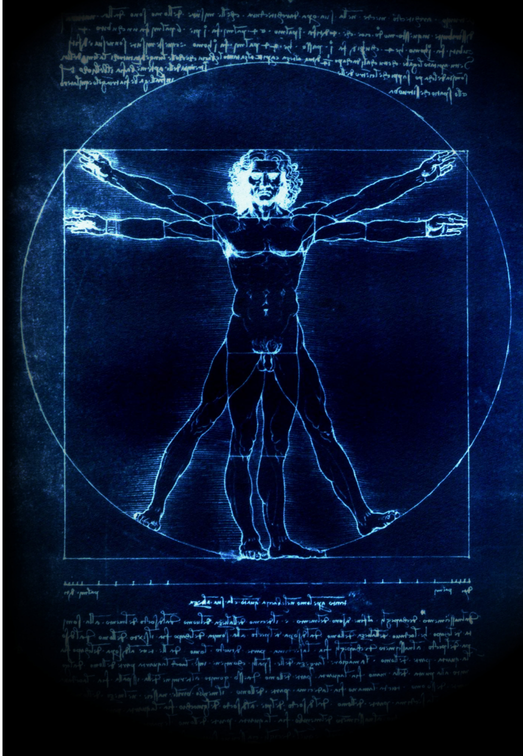
Can we have a statistical control on this process?

What role do models play?

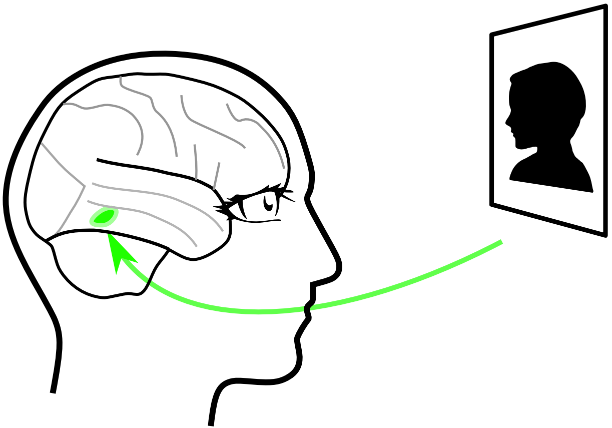
This talk

1. Evidence in (cognitive) neuroscience

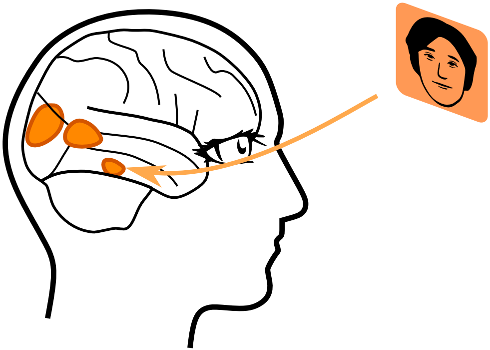
2. Models are overrated



1 Evidence in (cognitive) neuroscience

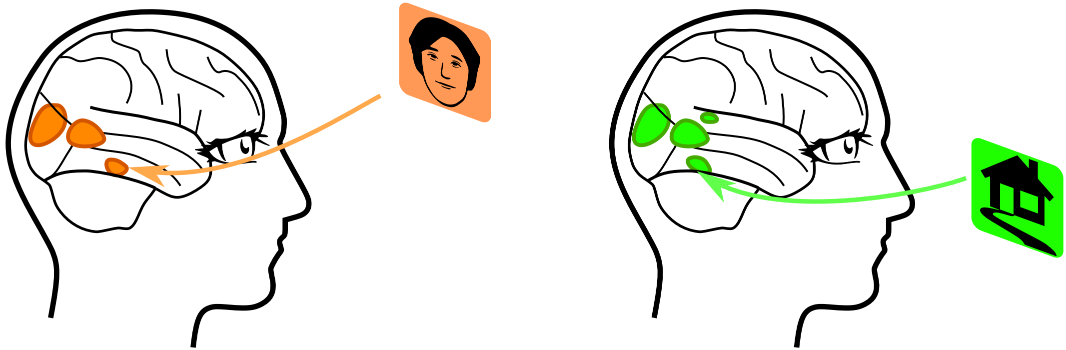


Probing a mental process via opposition



1 Craft an experimental condition that recruits it

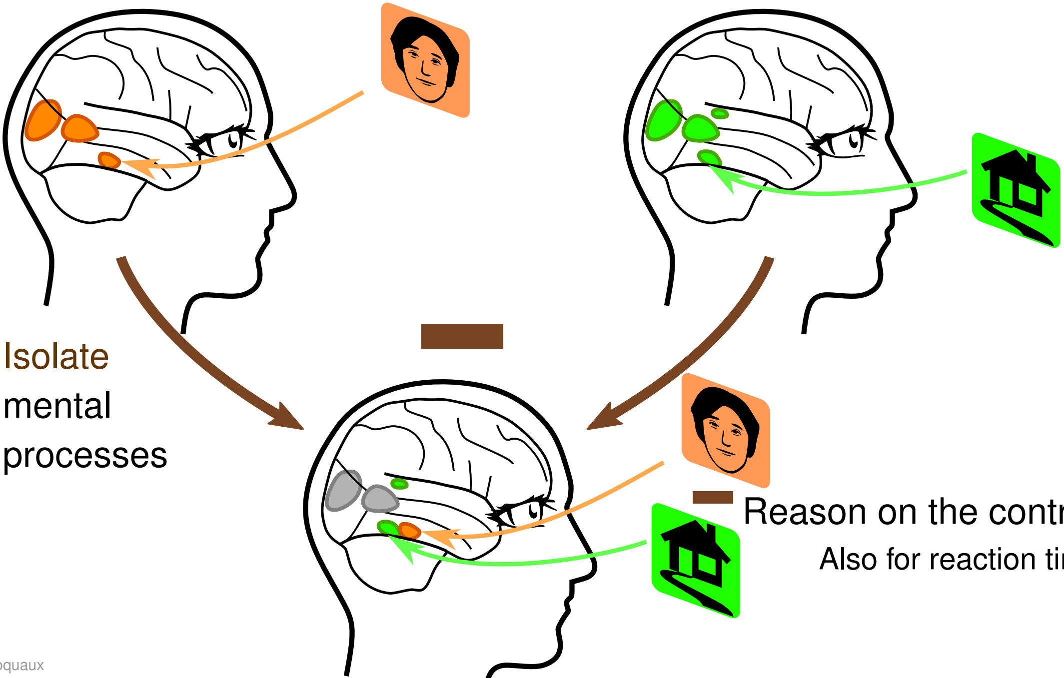
Probing a mental process via opposition



1 Craft an experimental condition that recruits it

2 Do an *elementary psychological manipulation*

Probing a mental process via opposition



Isolate mental processes

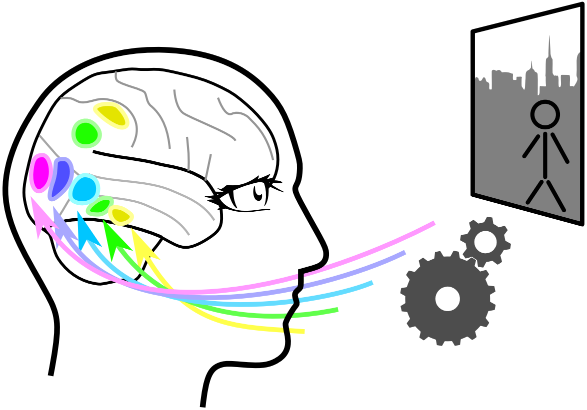
Reason on the contrast

Also for reaction times

The lens of the cognitive model

Psychological manipulations are designed and interpreted based on a cognitive model

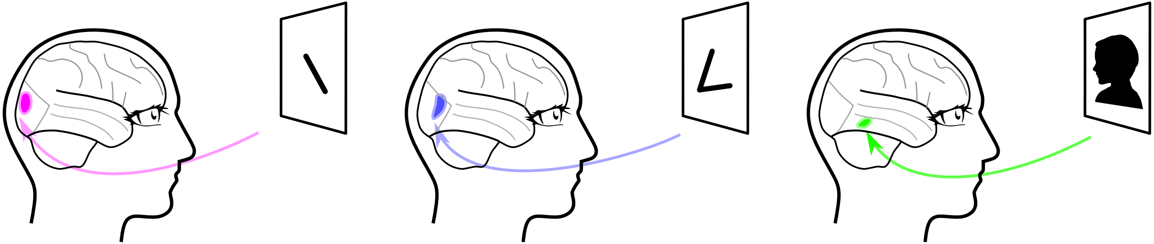
Experimental "paradigm"



- **Task & stimuli used** – should recruit the right mental processes
- **Opposition used** – should cancel out "nuisances"

The visual system: a paradigmatic example

- Successive experiments have revealed specialized regions

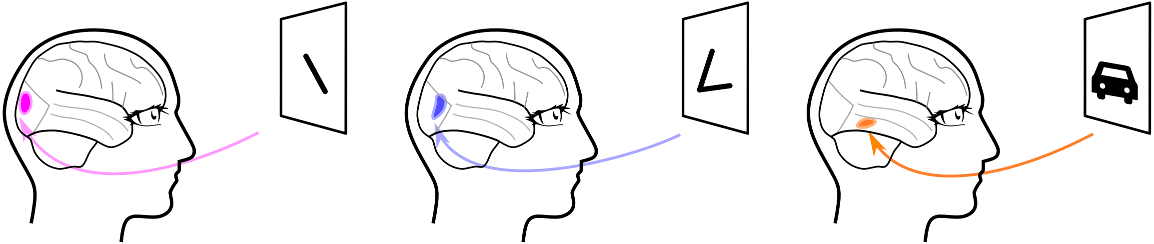


[Hubel and Wiesel 1959, Logothetis... 1995, Kanwisher... 1997]

The visual system: a paradigmatic example

- Successive experiments have revealed specialized regions

- But evidence is tied to a theory decomposing mental processes
  Is there a car area?



[Poldrack 2010]

Problem: Brain signals would struggle to debunk false theories

- Successive experiments have revealed specialized regions

- But evidence is tied to a theory decomposing mental processes
  Is there a car area?

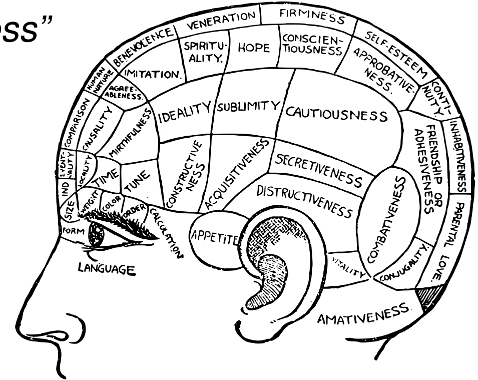
- Ingredients now considered invalid would yield significant differences

“philoprogenitiveness”

“alimentiveness”

“mirthfulness”

...

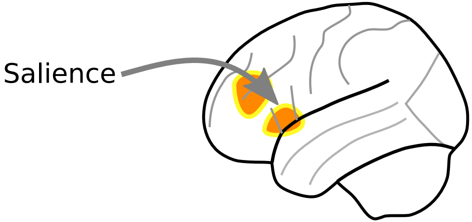


[Poldrack 2010]

Problem: The inference is the wrong way

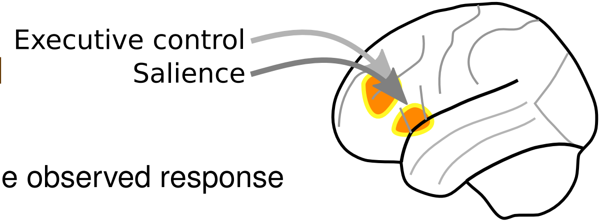
[Poldrack 2006]

What mental process is supported by this brain structure?



Problem: The inference is the wrong way

[Poldrack 2006]

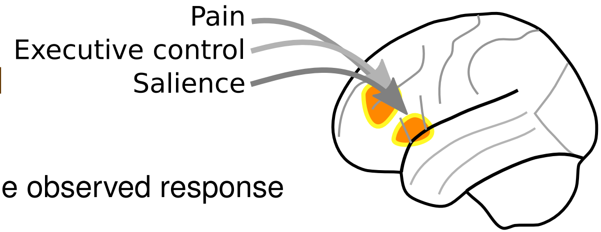


What mental process is supported by this brain structure?

The experimental manipulation *implies* the observed response

Problem: The inference is the wrong way

[Poldrack 2006]



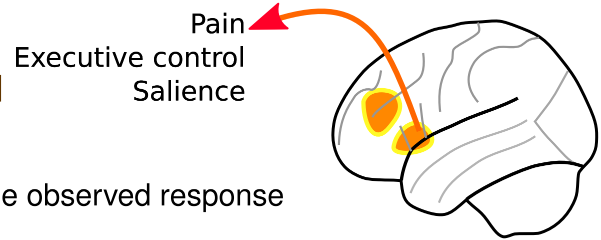
What mental process is supported by this brain structure?

The experimental manipulation *implies* the observed response

Empirical evidence: $\mathcal{P}(\text{neural activity}|\text{mental process})$

Problem: The inference is the wrong way

[Poldrack 2006]



What mental process is supported by this brain structure?

The experimental manipulation *implies* the observed response

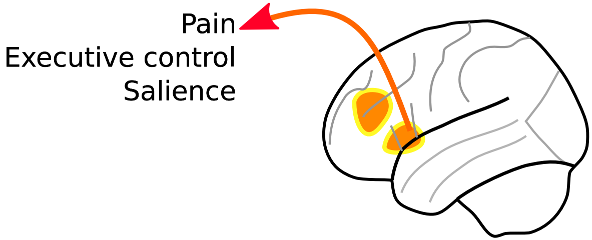
Empirical evidence: $\mathcal{P}(\text{neural activity}|\text{mental process})$

To conclude that **neural activity** $\Rightarrow$ **mental process**

- High-dimensional statistics (many brain regions / neurons)
- Requires data on many / all mental processes
- Ideally would be a causal claim

New methodology: predicting the task

Machine learning to predict
mental processes from activity



- High-dimensional statistics
  Machine learning: abandonning well-posed maximum likelihood

- Requires data on many / all mental processes
  Challenge = calibrated labeling of mental processes in tasks
  (not only oppositions)

- Ideally would be a causal claim — Let me come back to this

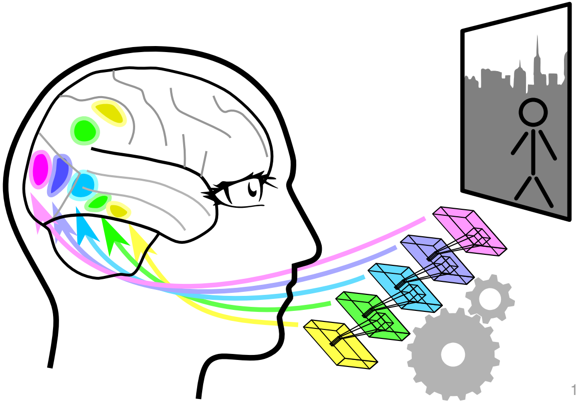
[Poldrack 2011, Varoquaux... 2018, Menuet... 2022]

New methodology: AI models for less reductionist task decomposition

Computer vision as a model for human vision

- Internal representations capture <u>all</u> aspects of natural stimuli
- Mapping them to brain responses with high-dimensional predictors

[Yamins... 2014]



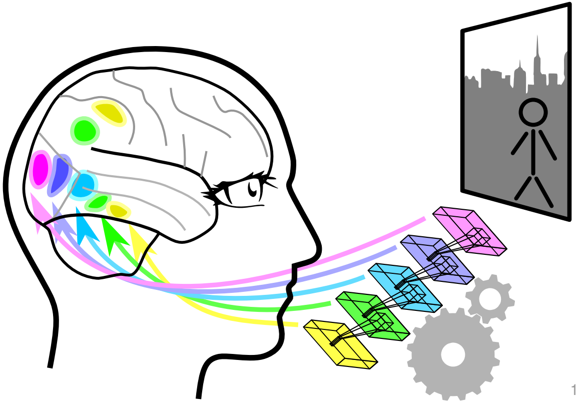
New methodology: AI models for less reductionist task decomposition

Computer vision as a model for human vision

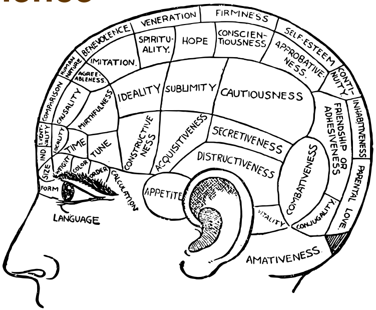
- Internal representations capture all aspects of natural stimuli
- Mapping them to brain responses with high-dimensional predictors

- Avoids choosing few ingredients/facets of a cognitive process (excess reductionism)
  [Varoquaux and Poldrack 2019]
- Can generalize across experimental paradigms
  [Eickenberg... 2017]

[Yamins... 2014]

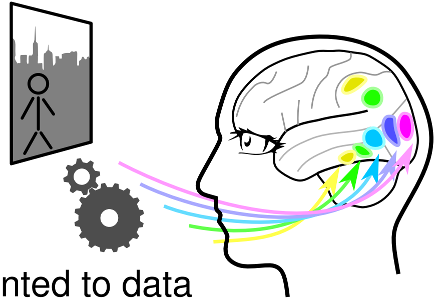


Evidence in cognitive neuroscience



- Focus on *significance* rather than *signal fit* leaves open doors to wrong models
- Well-posed models must be overly simple, and cannot answer the questions of interest

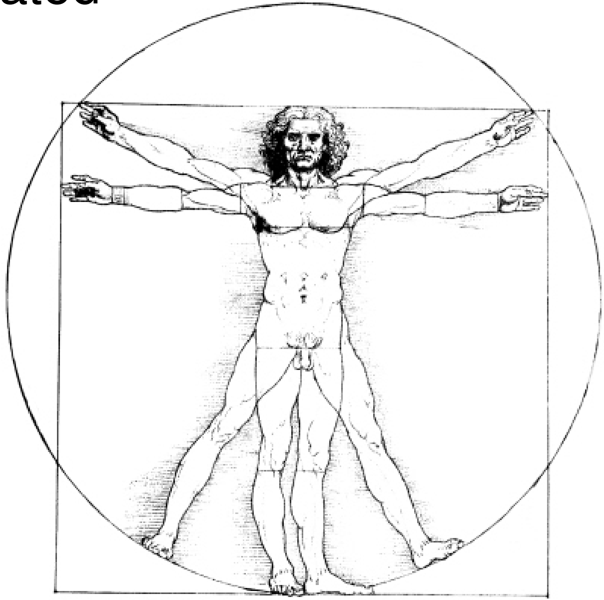
Machine learning / IA enables to model the complexity of the actual situations



But we want *understanding*

The answer does not lie in simplistic mechanistic models wich cannot be confronted to data

2 Models are overrated



Scientific progress and statistical evidence

Dominant framework of statistical reasoning:

- Formulating a probabilistic model from mechanical hypotheses
- Integrating empirical evidence (data) by fitting this model
- Reasoning from model parameters

Rigour breaks down with wrong modeling ingredients

Science needs more reasoning from model outputs

- For statistics: robustness to mis-specification
- Generalization grounds scientific theories

Black-box phenomenological data models are good for science

Teachings from history of science

Current view of physics, maths, chemistry...

Building models from the right ingredients – “first principles”

The past

Refining relevant constructs
from wrong models

The birth of mechanics

Early scientists (*eg* ancient Greece)

"natural motion of objects", no notion of force, or acceleration.

Observation of planetary motion (*eg* Kepler)

Search for regularities in planets – "harmonies"

The period squared is proportional to the cube of the major diameter of the orbit

Modern laws of dynamics (Newton)

Differential calculus ⇒ laws with force and acceleration

Unite observations of celestial and earthly motions

The birth of mechanics

Early scientists (*eg* ancient Greece)

"natural motion of objects", no notion of force, or acceleration.

Lacking key ingredients

Observation of planetary motion (*eg* Kepler)

Search for regularities in planets – "harmonies"

The period squared is proportional to the cube of the major diameter of the orbit

Phenomenological model[1] crucial

Modern laws of dynamics (Newton)

Differential calculus ⇒ laws with force and acceleration

Unite observations of celestial and earthly motions

Validity established by strong generalizability

Modern physics knows its laws?

Vulcan: false discovery of a planet (19th century)

Anomaly in Mercury's orbit not explained by Newtonian physics

$\Rightarrow$ invent and "observe" an additional planet, Vulcan

Theory laden observations

Modern physics knows its laws?

Vulcan: false discovery of a planet (19th century)

Anomaly in Mercury's orbit not explained by Newtonian physics

$\Rightarrow$ invent and "observe" an additional planet, Vulcan

Theory laden observations

Particle physics builds evidence with machine learning (today)

Fundamental laws of the universe = most precise theory ever

Particle detection by discriminating physics model with non-parametric background

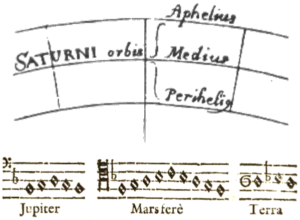
"Pure" models insufficient for "dirty" reality

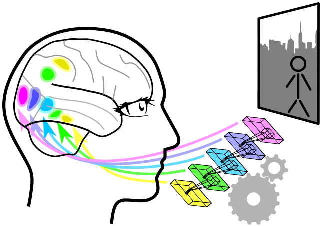
Phenomenological data fits have been crucial to science

- Science uses false models as means for truer theory
  [Wimsatt 2007]

- The reductionist aesthetics of “pure” simple mathematical theories is not adapted to the messy world beyond pure physics

- Generalization or prediction failures make or break scientific theories





Statistics and scientific evidence

- Validity
- Reasoning

  = more than formal problems

Validity of scientific findings – much more than statistical validity

External validity [Cook and Campbell 1979]

External validity asserts that findings apply beyond the study

Generalizability

Validity of scientific findings – much more than statistical validity

External validity [Cook and Campbell 1979]

External validity asserts that findings apply beyond the study

Generalizability

Constructs and their validity [Cronbach and Meehl 1955]

- Construct = abstract ingredients such as “intelligence”
- Construct validity: measures and manipulations actually capture the theoretical construct

Validity of scientific findings – much more than statistical validity

External validity [Cook and Campbell 1979]

External validity asserts that findings apply beyond the study

Generalizability

Constructs and their validity [Cronbach and Meehl 1955]

- Construct = abstract ingredients such as "intelligence"
- Construct validity: measures and manipulations actually capture the theoretical construct

Implicit realistic stances in theories

Realism = objective and mind-independent unobservable entities

Is intelligence a valid construct? How about a center of gravity?

Places implicit preferences on models beyond empirical evidence

Reasoning with statistical tools

Model reasoning [Cox 2006]

- Carefully craft a probabilistic model of the data
- Estimated model parameters are interpreted within its logic
  "data descriptions that are potentially causal" [Cox 2001]

Warranted reasoning [Baiocchi and Rodu 2021]

- Relies on warrants in the experiment (*eg* randomization)

Output reasoning [Breiman 2001, Baiocchi and Rodu 2021]

- Relies on capacity to approximate relations

Benefits of reasoning on outputs rather than models

Science needs black-box output reasoning

For statistical validity

Even expert modeling choices explore meaningful variability

- Model reasoning is conditional to the model
  parameters have a meaning in a model

- Imperfect science: 70 different teams of brain-imaging experts
  qualitatively different neuroscience findings [Botvinik-Nezer... 2020]

<u>Analytical variability breaks statistical control</u>

Output reasoning: milder conditions for statistical control

- Theoretical results in mispecified settings [Hsu... 2014]
- Multi-colinearity no longer an issue
- Higher-dimensional settings

$\Rightarrow$ Forces less reductionist choices

For broader scientific validity of findings

The only strong evidence is strong generalization

Model reasoning favors internal validity

Model reasoning often need “pure” models with little generalization

Fields without a unifying quantitative theory
tackle empirical evidence with overly reductionist lenses

Machine learning/AI can model the full problem space
and give testable generalization

For broader scientific validity of findings

The only strong evidence is strong generalization

Model reasoning favors internal validity

Understanding and reasoning without parametric models

Counterfactual reasoning, causal inference
with machine-learning models

- Causality = intervention
- Observational data reflects uncontroled interventions
(hospital visit when sick)
- Causal effects = interventions with same context (*eg* health status)
- Causal inference with machine learning:
predicting outcome from intervention and context

[Rose and Rizopoulos 2020, Angrist and Pischke 2008]

AI gives statistical methods for imperfect theories

- Model reasoning has no guarantees for imperfect models

- Output reasoning relaxes modeling constraints

- Scientific roadblocks are on model ingredients, not functional forms

Proposal

- Gauge models more on their predictions than their ingredients
- Develop scientific methods around model predictions

counterfactual reasoning, model comparison, feature importances

References I

J. D. Angrist and J.-S. Pischke. *Mostly harmless econometrics*. Princeton university press, 2008.

M. Baiocchi and J. Rodu. Reasoning using data: Two old ways and one new. *Observational Studies*, 7(1):3–12, 2021.

R. Botvinik-Nezer, F. Holzmeister, C. F. Camerer, A. Dreber, J. Huber, M. Johannesson, M. Kirchler, R. Iwanir, J. A. Mumford, R. A. Adcock, ... Variability in the analysis of a single neuroimaging dataset by many teams. *Nature*, 582(7810):84–88, 2020.

L. Breiman. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science*, 16(3):199–231, 2001.

T. Cook and D. Campbell. *Quasi-experimentation: Design and analysis issues for field settings 1979 Boston*. MA Houghton Mifflin, 1979.

D. R. Cox. [statistical modeling: The two cultures]: Comment. *Statistical science*, 16(3): 216–218, 2001.

D. R. Cox. *Principles of statistical inference*. Cambridge university press, 2006.

References II

L. J. Cronbach and P. E. Meehl. Construct validity in psychological tests. *Psychological Bulletin*, 52:281, 1955.

M. Eickenberg, A. Gramfort, G. Varoquaux, and B. Thirion. Seeing it all: Convolutional network layers map the function of the human visual system. *NeuroImage*, 152:184–194, 2017.

D. Hsu, S. Kakade, and T. Zhang. Random design analysis of ridge regression. *Foundations of Computational Mathematics*, 14, 2014.

D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.*, 148:574–591, 1959.

N. Kanwisher, J. McDermott, and M. M. Chun. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.*, 17(11):4302–4311, 1997.

N. K. Logothetis, J. Pauls, and T. Poggio. Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, 5:552, 1995.

R. Menuet, R. Meudec, J. Dockès, G. Varoquaux, and B. Thirion. Comprehensive decoding mental processes from web repositories of functional brain images. *Scientific Reports*, 12 (1):1–14, 2022.

References III

R. Poldrack. Can cognitive processes be inferred from neuroimaging data? *Trends in cognitive sciences*, 10:59, 2006.

R. A. Poldrack. Mapping mental function to brain structure: how can cognitive neuroimaging succeed? *Perspectives on psychological science*, 5:753, 2010.

R. A. Poldrack. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron*, 72:692, 2011. This review show how decoding can be used on large-scale databases to ground formal reverse inference, capturing how selectively a brain area is activated by a mental process.

S. Rose and D. Rizopoulos. Machine learning for causal inference in biostatistics. *Biostatistics*, 21(2):336–338, 2020.

G. Varoquaux. Ai as statistical methods for imperfect theories. In *NeurIPS 2021 AI for Science Workshop*, 2021.

G. Varoquaux and R. A. Poldrack. Predictive models avoid excessive reductionism in cognitive neuroimaging. *Current opinion in neurobiology*, 55:1–6, 2019.

References IV

G. Varoquaux, Y. Schwartz, R. A. Poldrack, B. Gauthier, D. Bzdok, J.-B. Poline, and B. Thirion. Atlases of cognition with large-scale human brain mapping. *PLOS Computational Biology*, 14(11):1–18, 11 2018.

W. C. Wimsatt. *Re-engineering philosophy for limited beings: Piecewise approximations to reality*. Harvard University Press, 2007.

D. L. Yamins, H. Hong, C. F. Cadieu, E. A. Solomon, D. Seibert, and J. J. DiCarlo. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc Natl Acad Sci*, page 201403112, 2014. This study shows that models of neural response based on computer-vision artificial networks explain brain activity better than classic theoretical-neuroscience models of vision.