Data mining and modeling for behavioral sciences and beyond

Valentin Wyart professor
Lucas Benjamin teaching assistant

Lab. de Neurosciences Cognitives et Computationnelles (LNC^2)
Institut National de la Santé et de la Recherche Médicale
Ecole Normale Supérieure, Université PSL

valentin.wyart@ens.psl.eu

PSL UNIVERSITÉ PARIS

Data Science Program (certifying minor at PSL)

https://psl.eu/en/programmes-gradues/programme-data

PR[AI]RIE

PaRis Artificial Intelligence Research InstitutE

Paris Artificial Intelligence Research Institute

https://prairie-institute.fr

Data mining and modeling

- We are currently facing an explosion of data across domains and disciplines.
- The ability to manipulate and understand large amounts of complex, multidimensional data has become critical in science (and for many applications outside academia).
- Can you give a few examples?

Data mining and modeling

- Data mining: manipulate data
  Afternoon classes = practical sessions

- Data modeling: understand data
  Morning classes = lectures

- Do you have a working Python environment?
  Come to me during the break if not.

Data mining and modeling

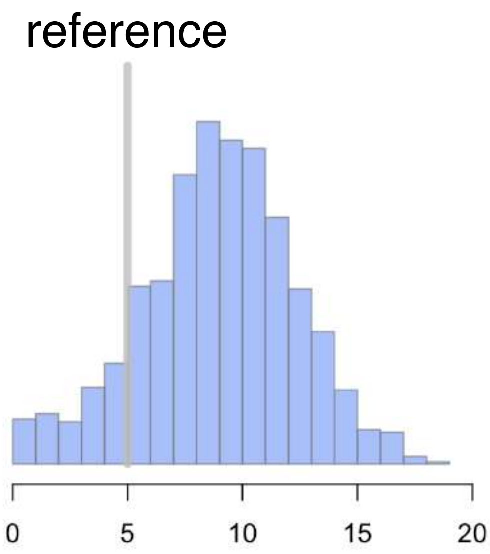
- Today's practical session: data statistics
  2.00pm, this room

- Au programme:
  - ✓ plot data
  - ✓ compute data metrics
    what metrics do you know?
    can the mean be a misleading metric? when?
    what is a probability density function?
  - ✓ identify correlations in data
    does correlation mean causation? why (not)?

Data mining and modeling

- Compare mean of data variable to reference
  Name of this test: 1-sample *t*-test

- Compare means of two data variables
  Name of this test: 2-sample *t*-test



Data mining and modeling

- Compare mean of data variable to reference
  Name of this test: 1-sample *t*-test
- Compare means of two data variables
  Name of this test: 2-sample *t*-test
- Why is it already a model of the data?
- Difference between a statistical model and a computational model of behavioral data

Data mining and modeling

- Uncertainty about data: what does it mean?
  Name factors that influence data uncertainty.
- How do we report uncertainty about the mean?
- Difference between a point estimate and a confidence interval
- Difference between analytical and empirical measures of data uncertainty
  Example: bootstrapping of uncertainty metrics

Data mining and modeling

- Data mining and modeling is not only about playing with data, but also thinking about data.
- Approaches in data mining and modeling can easily be misused, they can provide nonsensical answers, and you need to think about data to tell the difference.
- Introducing you to thinking about data is maybe the most important aim of this PSL-week.

Class evaluation

- Create 5 groups of 3 students each
- In-depth analysis of behavioral data collected in a slot machine game (two-armed bandit)
- Use data mining <u>and</u> modeling approaches seen during lectures <u>and</u> practical sessions
- <u>Objective:</u> identify the latent cognitive strategy that drives behavior (<u>different</u> for each group)
- Group presentation (15 min/group) on Friday

What is a model of cognition?

- Statistical models are used to summarize and describe behavioral data.
- Computational models are used to understand how behavioral data have been generated.
- This morning, we will discuss computational models of cognition = cognitive modeling.

What is a model of cognition?

- Cognitive modeling aims at understanding behavior in terms of its underlying cognitive processes.
- Cognitive modeling proceeds by building mathematical descriptions and computer algorithms of these processes that are able to reproduce the studied behavior.

What is a model of cognition?

- Not a statistical model of effect size: a *t*-test of response time differences between experimental conditions is not a cognitive model.
- Cognitive modeling aims at understanding not whether RTs differ between conditions, but rather why RTs differ between these two conditions.
- Cognitive models are computational instances of theories of cognition and behavior.

What is a model of cognition?

- Cognitive psychology typically proceeds by contrasting behavior between different experimental conditions.
- Cognitive modeling, like cognitive psychology, relies on contrastive methods to identify which features of these mathematical descriptions are necessary to reproduce behavior:
  - ✓ comparison between different models
  - ✓ simulation of different models

What is a model of cognition?

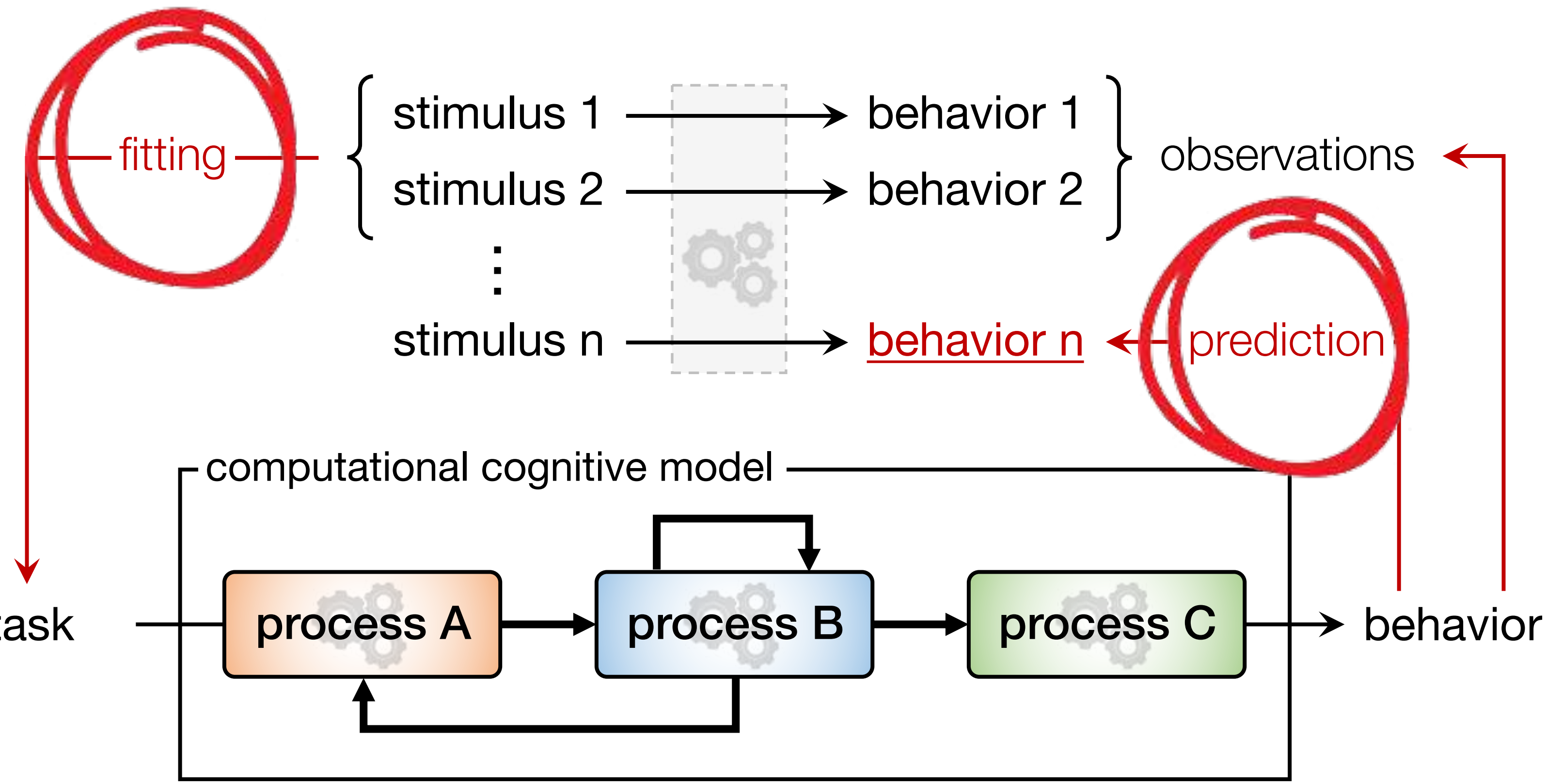
- Unlike cognitive psychology, cognitive modeling can be used to understand behavior in a single experimental condition.
- But like cognitive psychology, cognitive modeling heavily depends on experimental design to yield interpretable conclusions.

What is a model of cognition?

- Poor experimental design can result in different models reproducing the same behavior.
- Like cognitive psychology, cognitive modeling requires careful experimental design – and even several additional *a priori* analyses – to deliver its promises.

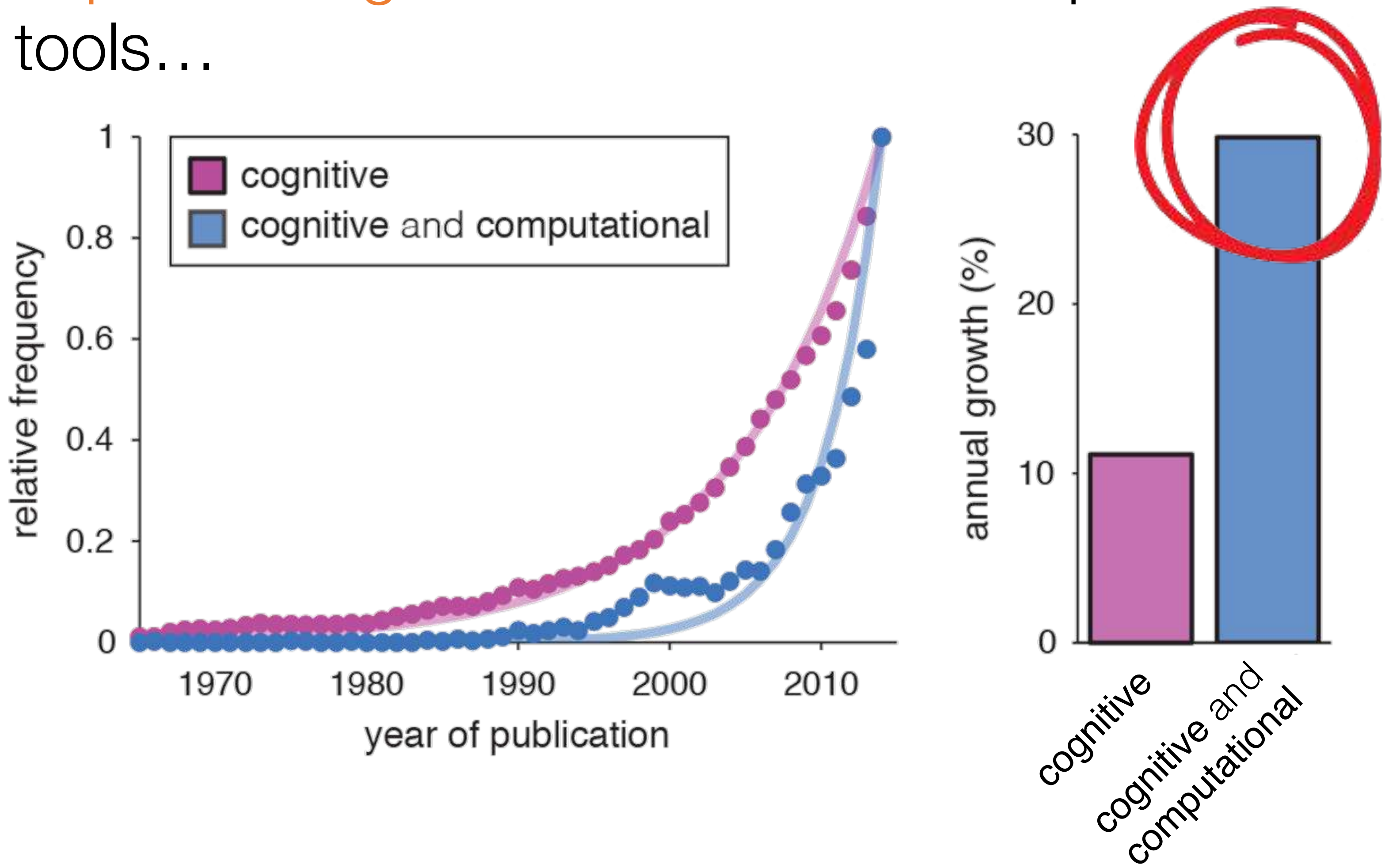
What is a model of cognition?

- Fitting vs predicting behavior



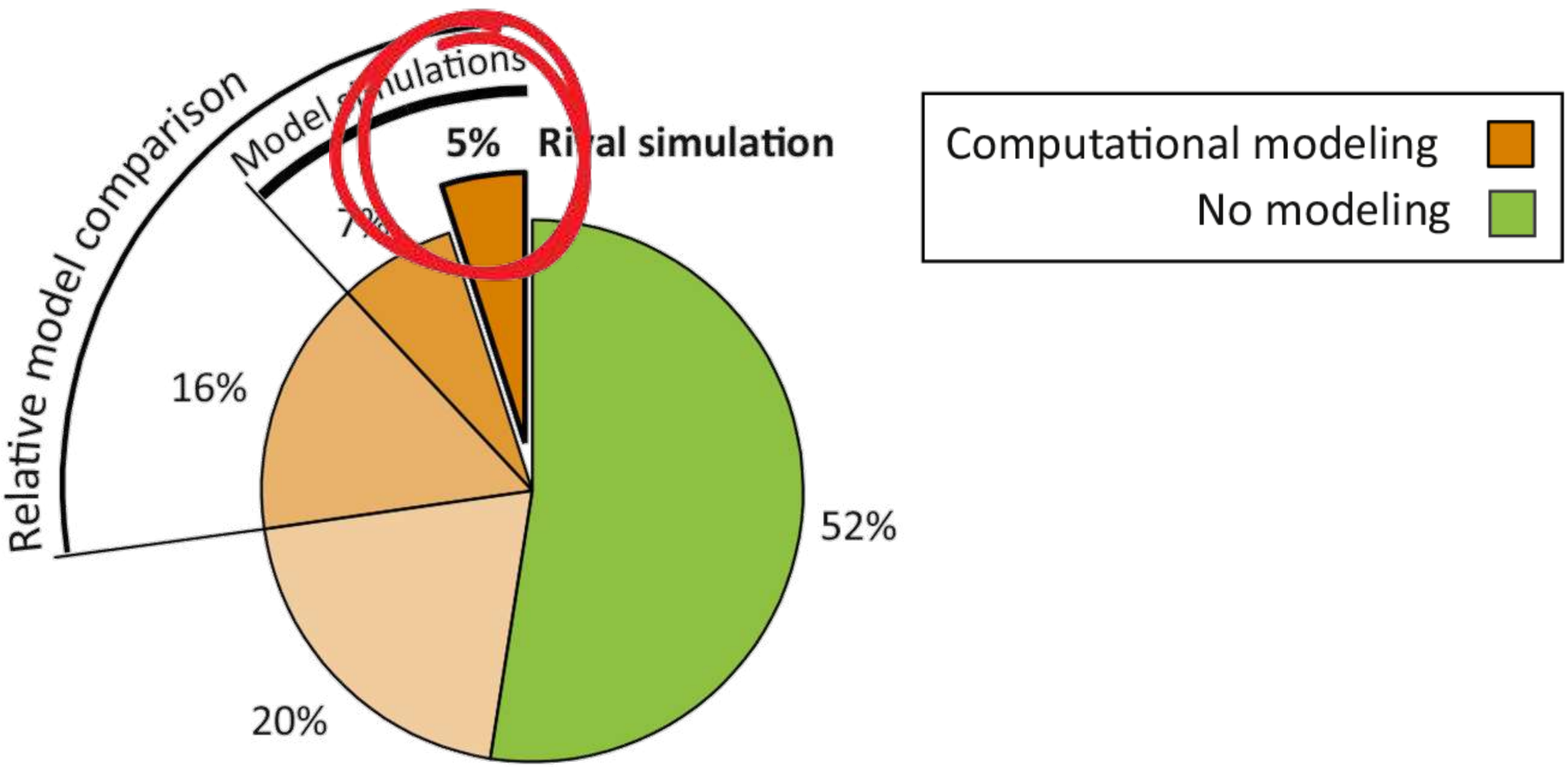
What is a model of cognition?

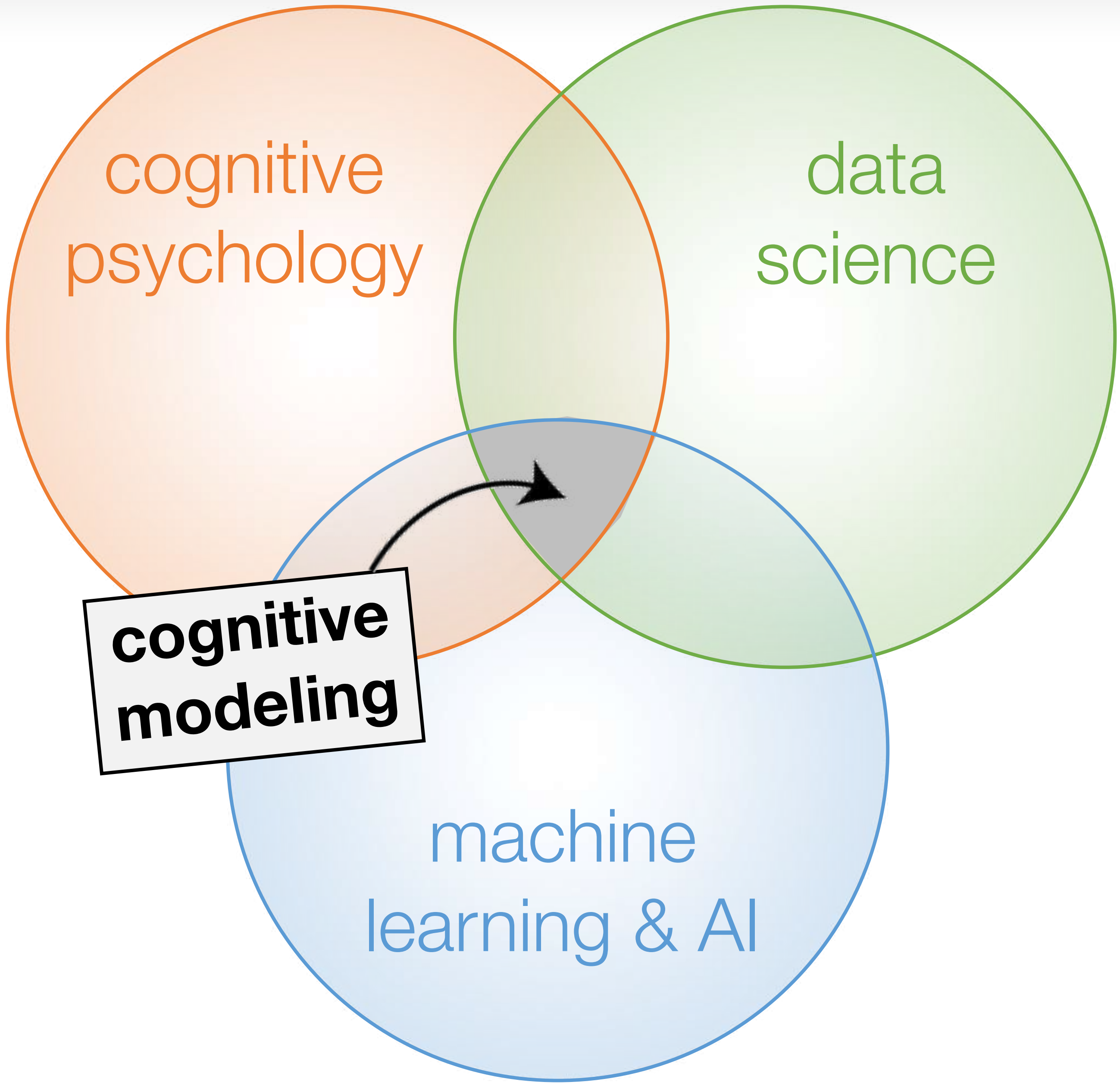
- Exponential growth in the use of computational tools…



What is a model of cognition?

- Exponential growth in the use of computational tools… mostly without valid conclusions?

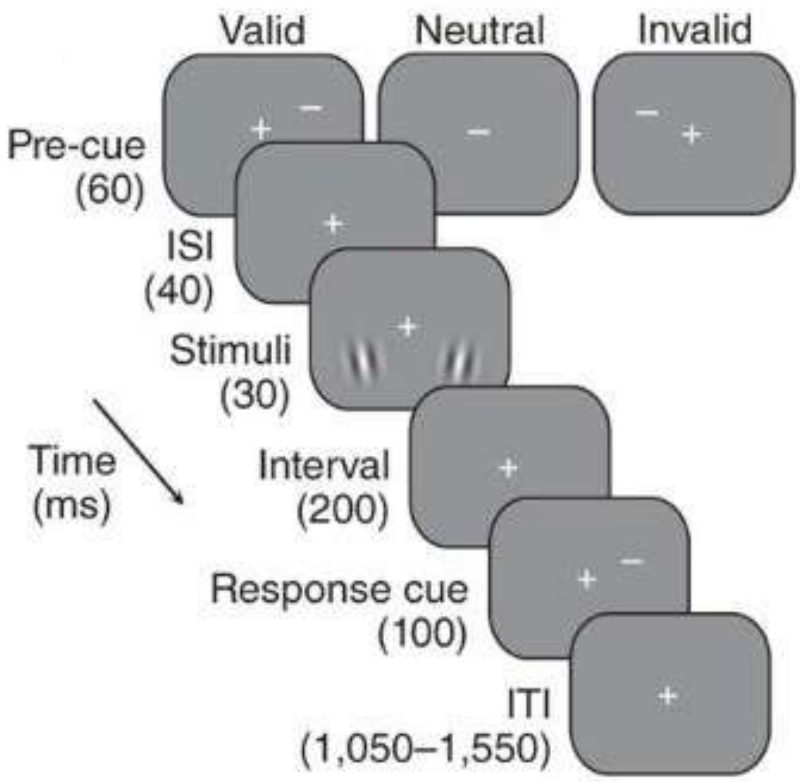



cognitive psychology
data science
machine learning & AI
cognitive modeling

What is a model of cognition?

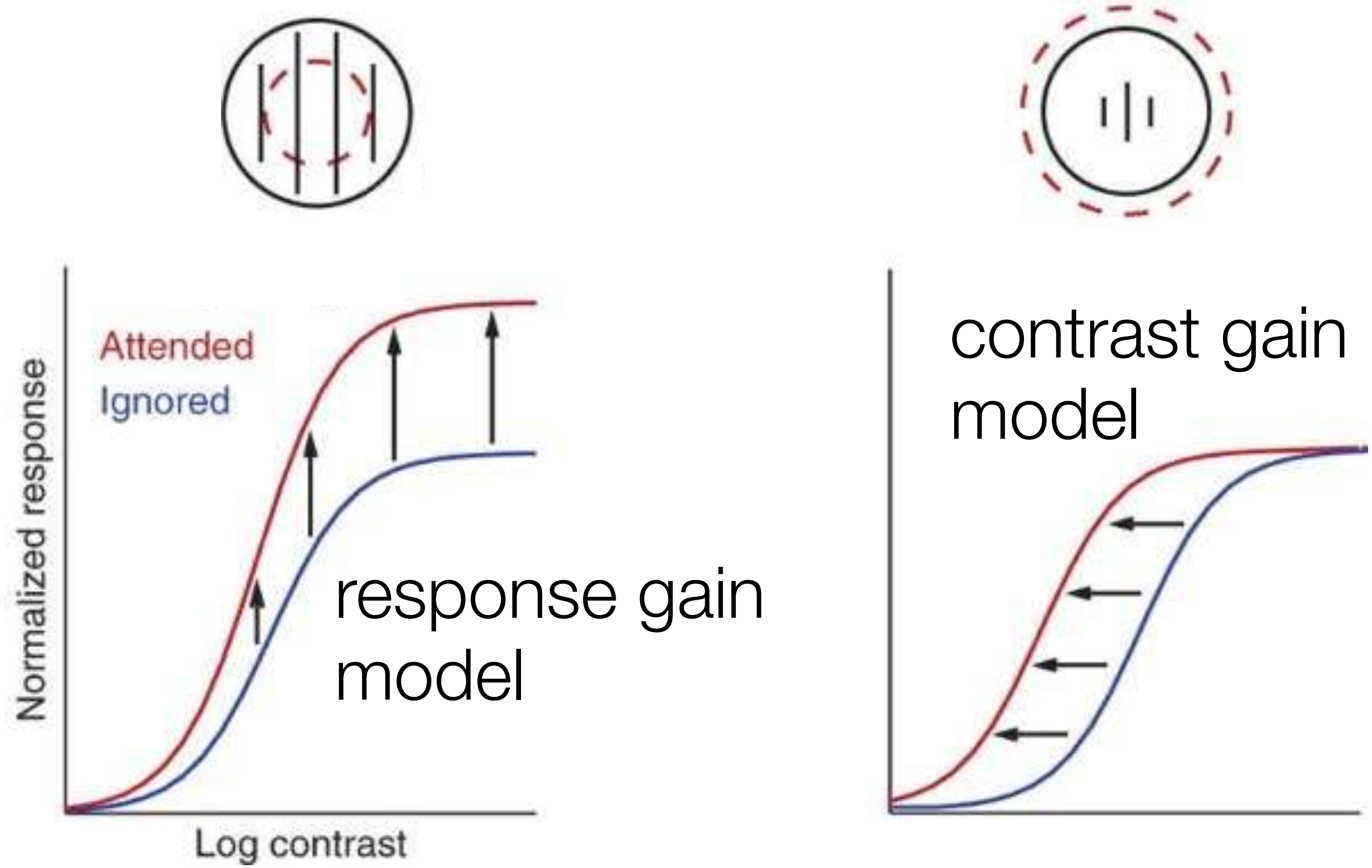
- Connections with cognitive psychology
  - ✓ shared goal:
    understand the human mind
  - ✓ shared techniques:
    design controlled experiments that target specific cognitive processes
- Cognitive psychology tests whether behavior differs between conditions, cognitive modeling aims at understanding why.

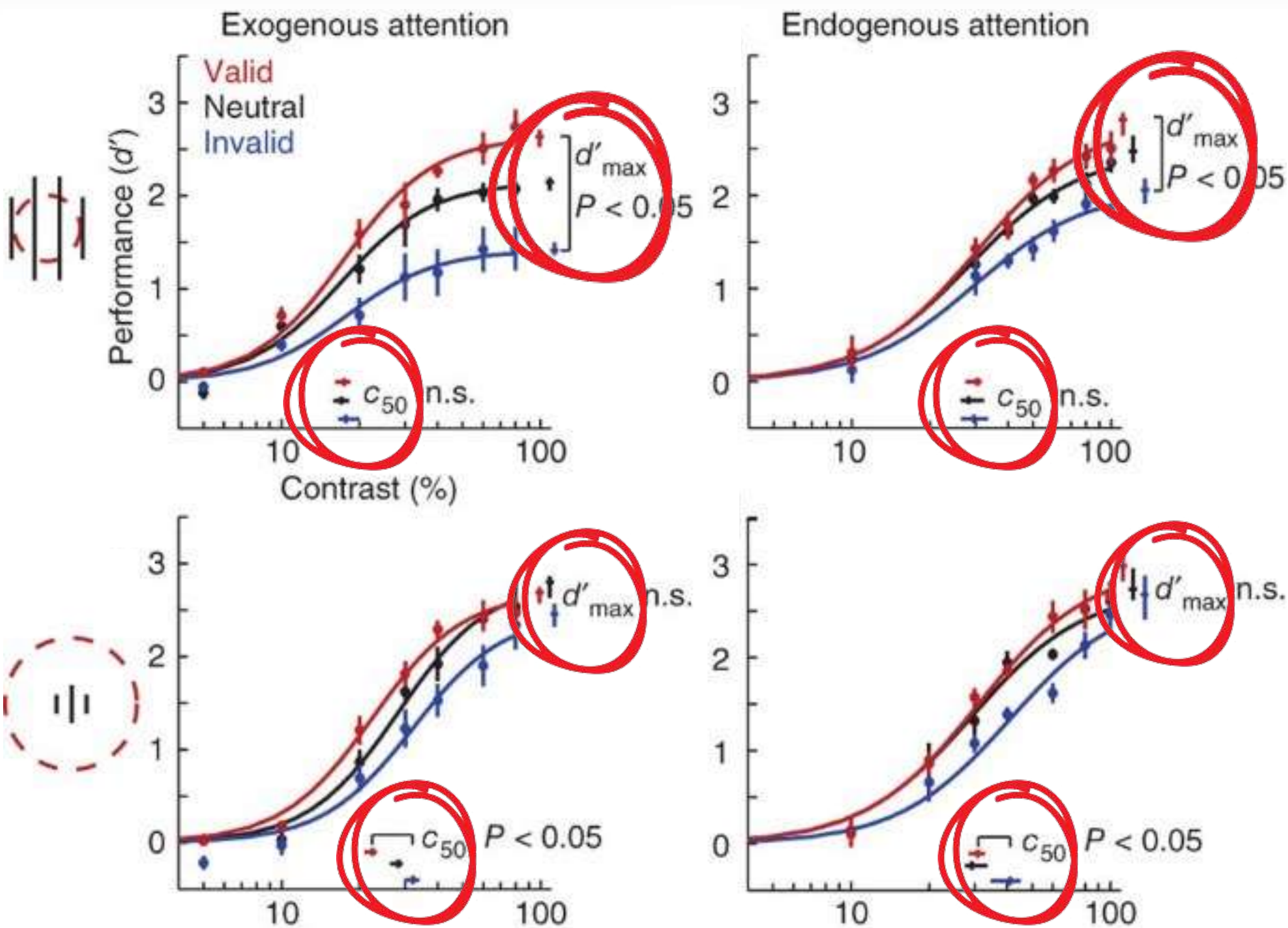
exogenous attention



large stimuli

small stimuli





Exogenous attention
Endogenous attention
Valid
Neutral
Invalid
Performance (d′)
Contrast (%)
d′max
P < 0.05
c50
n.s.
d′max
n.s.
c50
P < 0.05

What is a model of cognition?

- Cognitive psychology reveals significant differences between experimental conditions.
- Cognitive modeling provides computational mechanisms of attention that explain these differences between conditions.

Hold on a sec, cognitive psychology does offer explanations for differences between conditions!

Herrmann et al. (2010) When size matters: attention affects performance by contrast or response gain. *Nature Neuro.*

What is a model of cognition?

- Cognitive psychology uses rhetorical statements to explain differences between conditions.
- Such qualitative statements can be flawed by internal inconsistencies, logical contradictions, and theoretical weaknesses.
- A cognitive model is used as a quantitative proof of the internal coherence and completeness of the theory it is based upon.

What is a model of cognition?

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 8 (2007) 135–142

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Editorial

The cognitive modeling of human behavior: Why a model is (sometimes) better than 10,000 words

Action editor: Ron Sun

Danilo Fum [a,*], Fabio Del Missier [a], Andrea Stocco [b]

[a] *Università degli Studi di Trieste, Italy*
[b] *Carnegie Mellon University, United States*

Received 24 July 2007; accepted 24 July 2007
Available online 31 July 2007

1. Introduction

This special issue of Cognitive Systems Research presents a collection of remarkable papers on cognitive modeling based on communications delivered at ICCM-2006, the Seventh International Conference on Cognitive Modeling (Fum, Del Missier, & Stocco, 2006) held in Trieste, Italy, from April 5th to 8th, 2006. Being the organizers and chair-

allow not only to understand the how and why of the old things, but also to predict the happening of new ones.

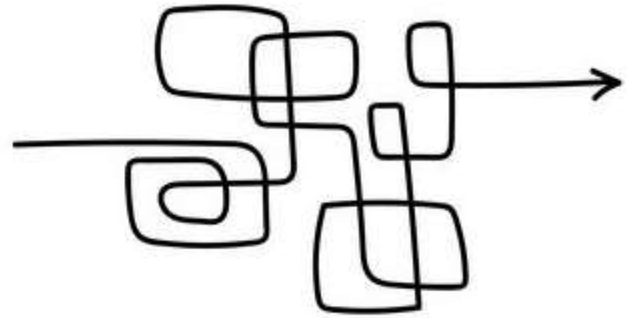
Within cognitive science we are trying to uncover how the mind works. Aiming toward this end, cognitive scientists have been developing an impressive array of empirical methods encompassing observational and correlational studies, human and animal experimentation, case studies of brain-damaged patients, physiological recordings and, more recently, neuroimaging techniques. Here we are inter-

What is a model of cognition?

- Connections with data science
  - ✓ shared goal:
    build computer algorithms to explain/predict behavioral data
  - ✓ shared techniques:
    formulate, simulate, fit, compare, validate computer algorithms against behavioral data
- Data science aims (first) at predicting data, cognitive models aim at understanding behavior.

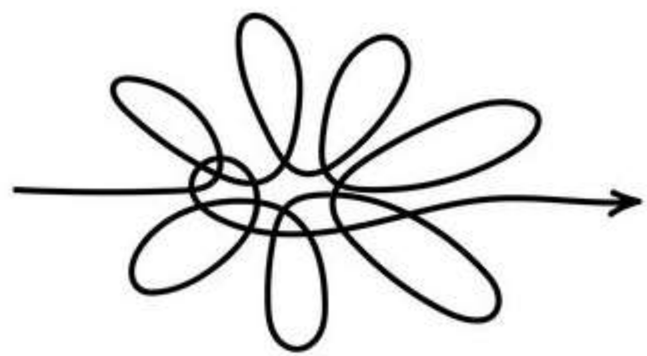
model A

model B

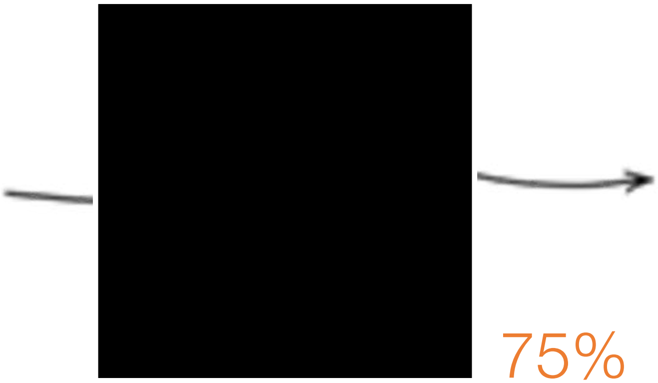


model C

model D



model A



75%

model B

78%

model C

68%

model D

72%

model A

75%

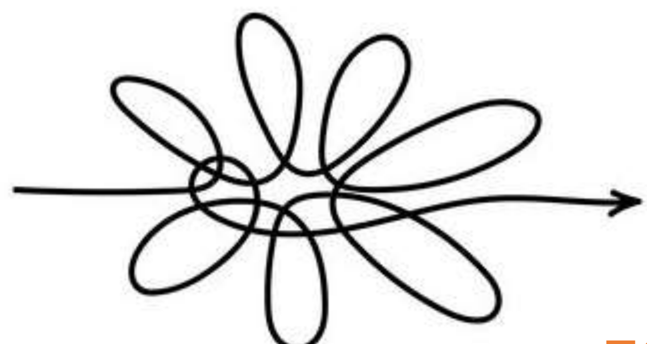
model B

78%

model C

68%

model D

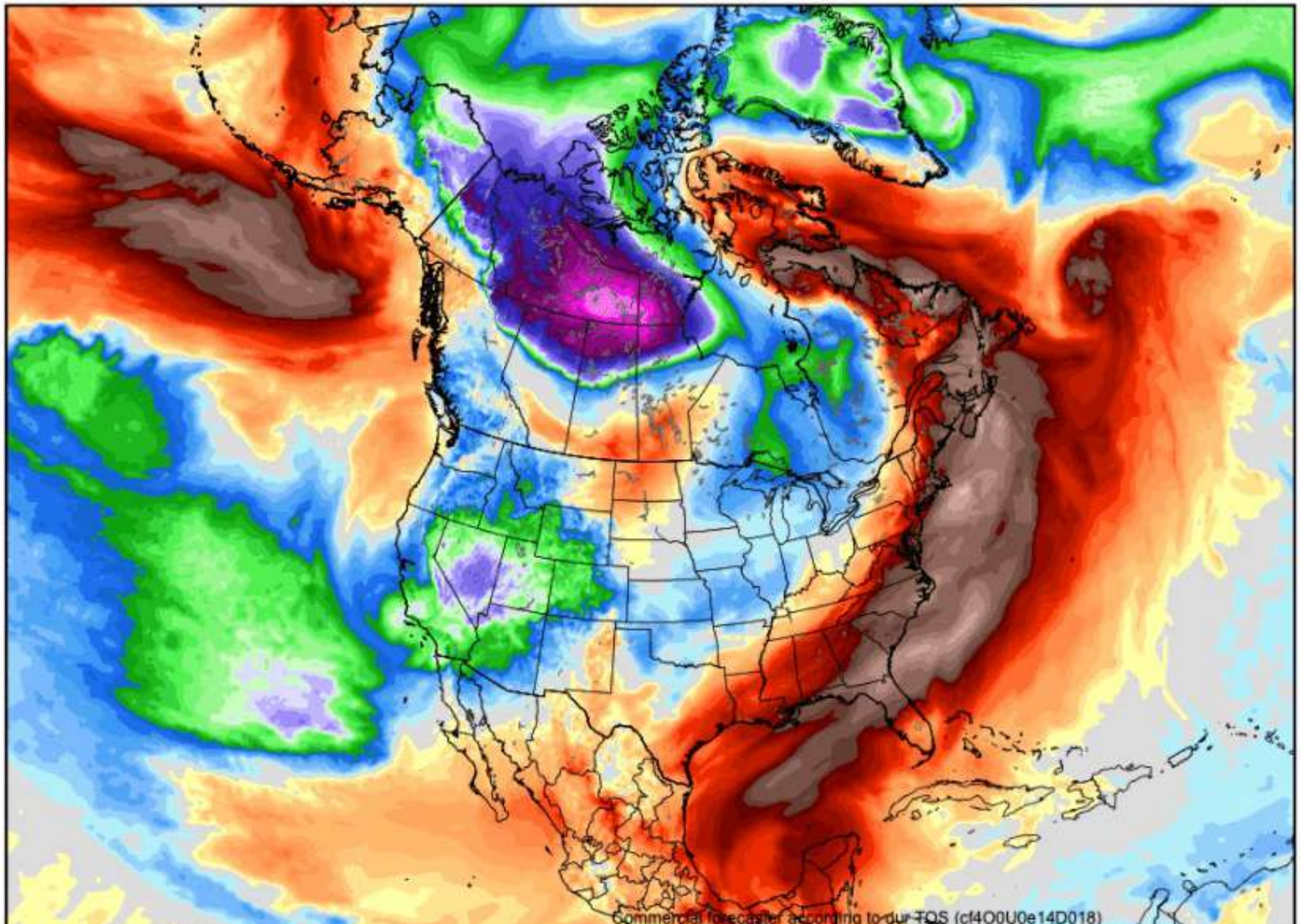


72%

What is a model of cognition?

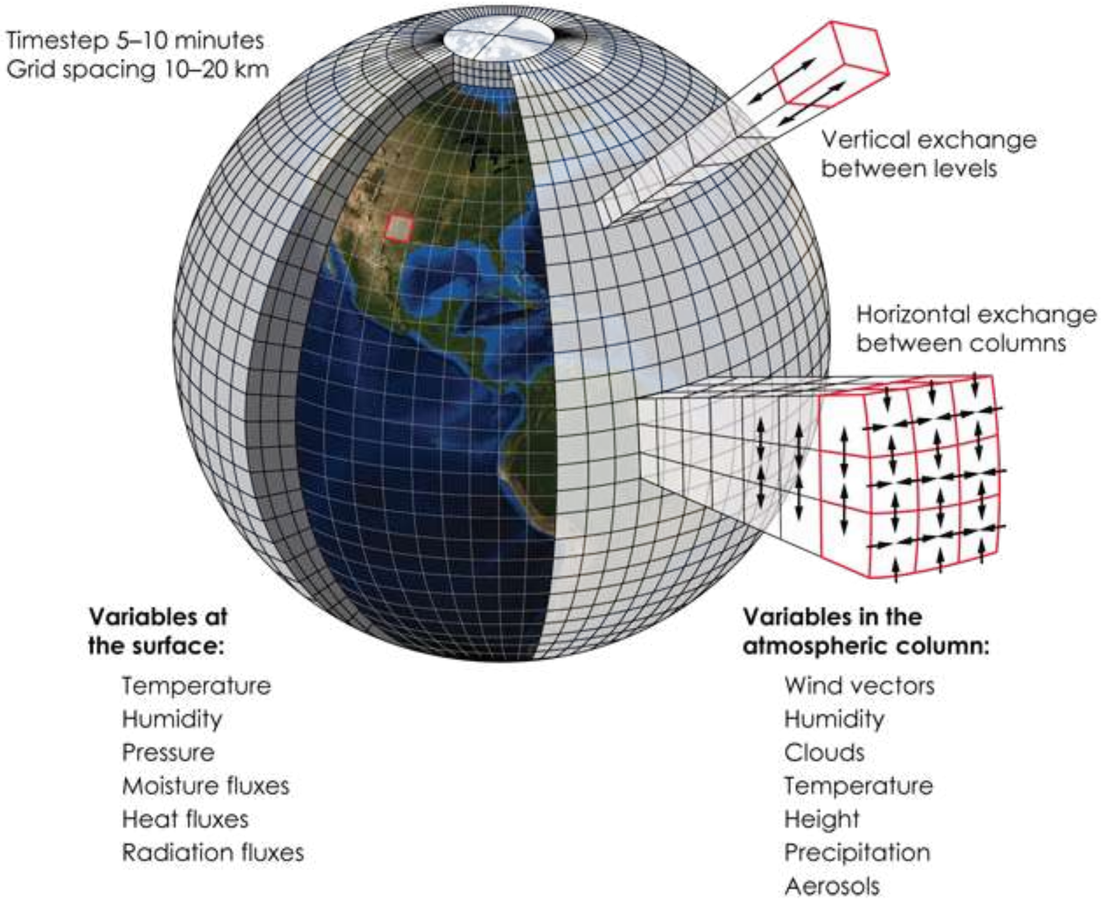
- Data science cares about predictive accuracy, not about the data generation process.
- Cognitive modeling cares about the underlying processes that have generated the data.

What is a model of cognition?



What is a model of cognition?

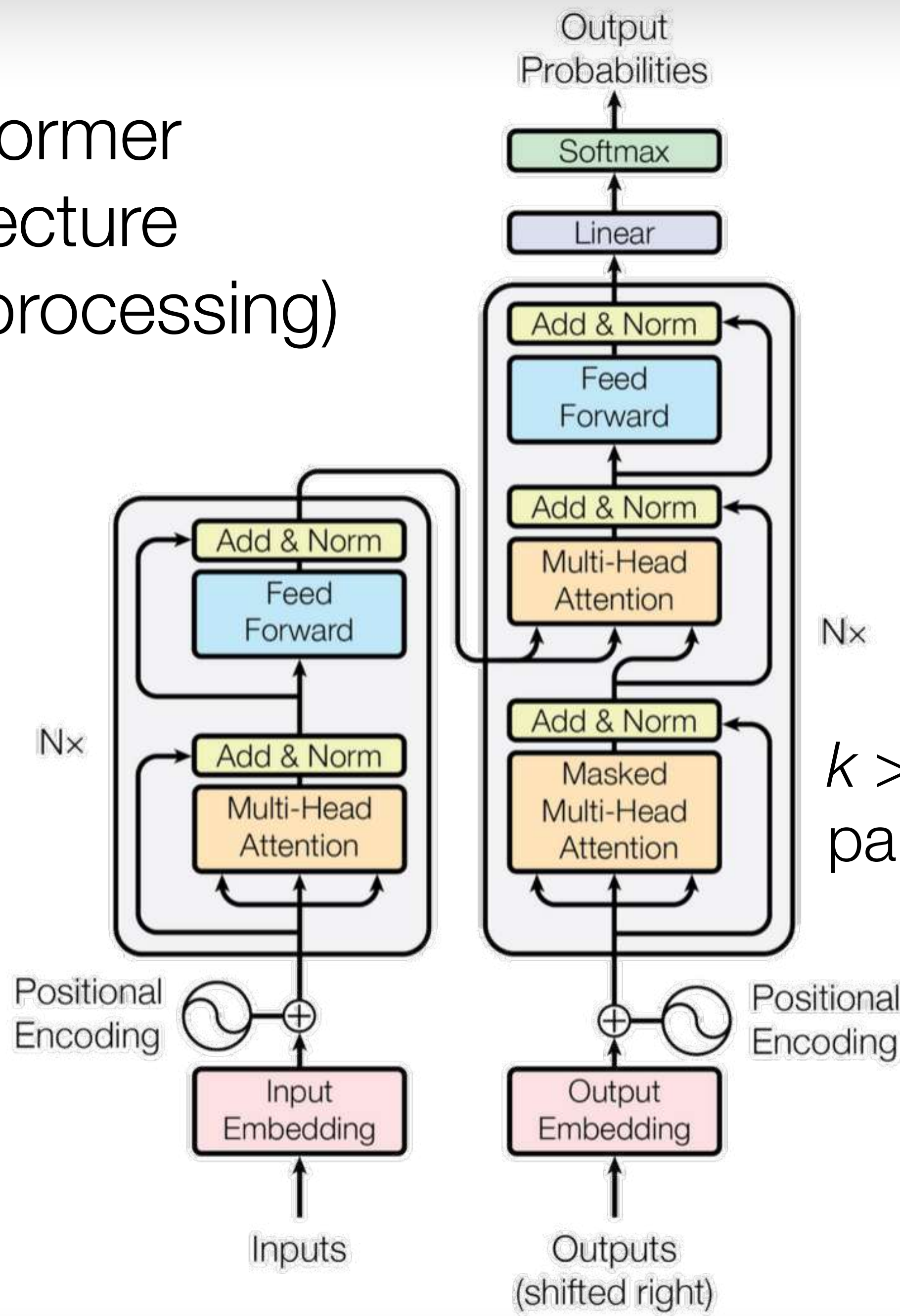
weather modeling



What is a model of cognition?

- Connections with machine learning & AI
  - ✓ shared goal:
    build models of the mind
  - ✓ shared techniques:
    reinforcement learning, particle filtering, pattern classification, neural networks…
- AI aims at maximizing performance, cognitive modeling aims at mimicking the human mind

Transformer architecture (language processing)

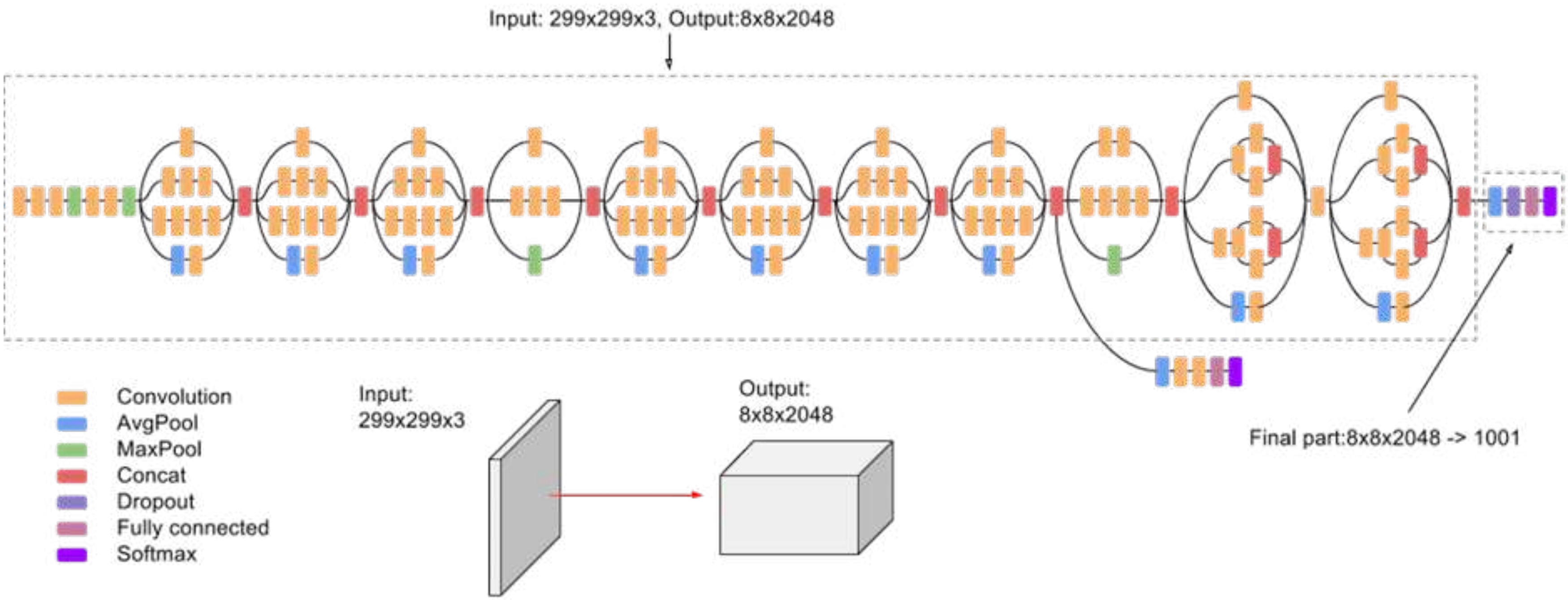


trained to predict next word/token

$k > 10^9$ free parameters

Inception v3 architecture (image recognition)

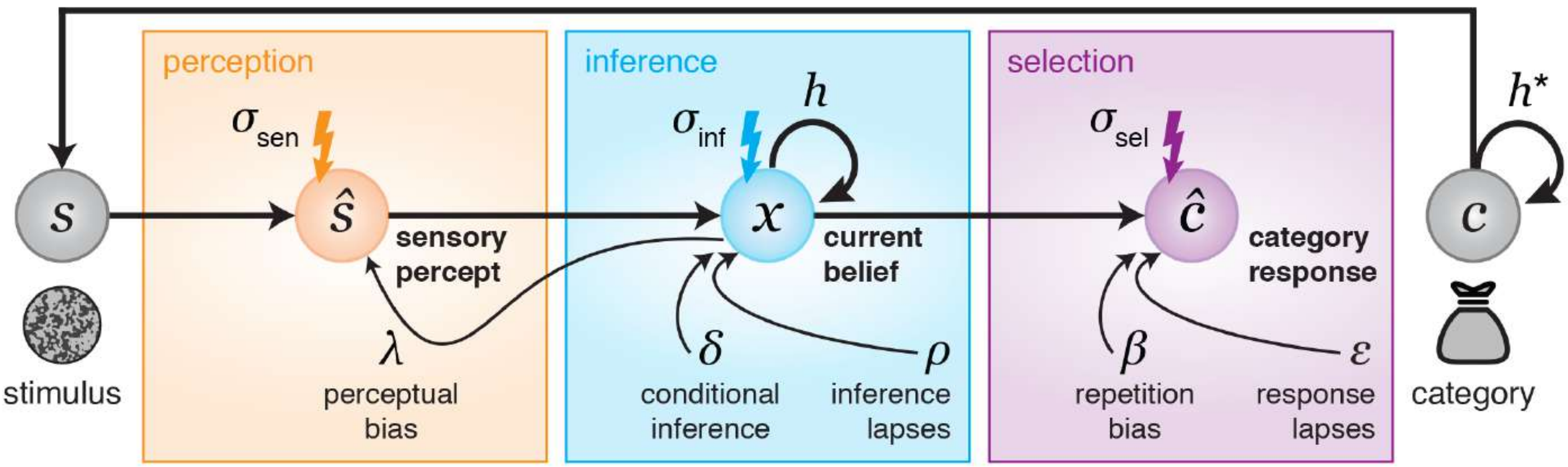
$k > 10^9$ free parameters



trained to maximize recognition accuracy

Hidden-state inference architecture (stimulus categorization)

$k = 8$ free parameters

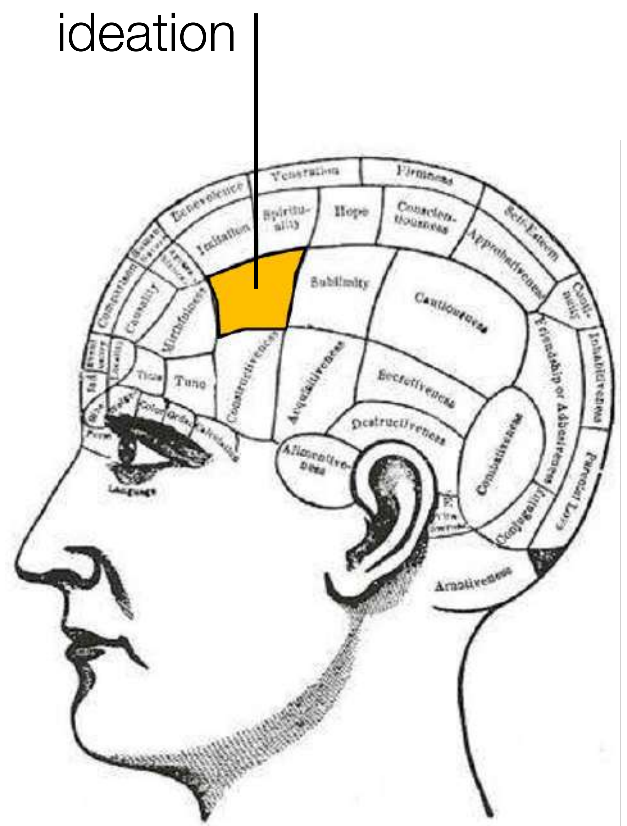


trained to reproduce human behavior

Why modeling cognition?
We can look into brains!

- Does mind = brain?

  From phrenology…
  FJ Gall (1758-1828)

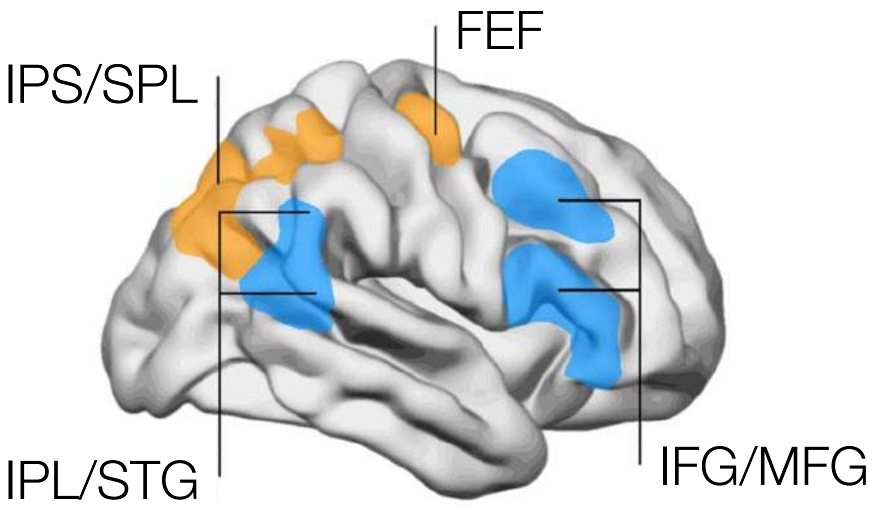


Why modeling cognition?
We can look into brains!

- Does mind = brain?

From phrenology…
FJ Gall (1758-1828)

to brain mapping
e.g., OHBM (next conference in June 2024 in Seoul)



Organization for Human Brain Mapping
Advancing Understanding of the Human Brain

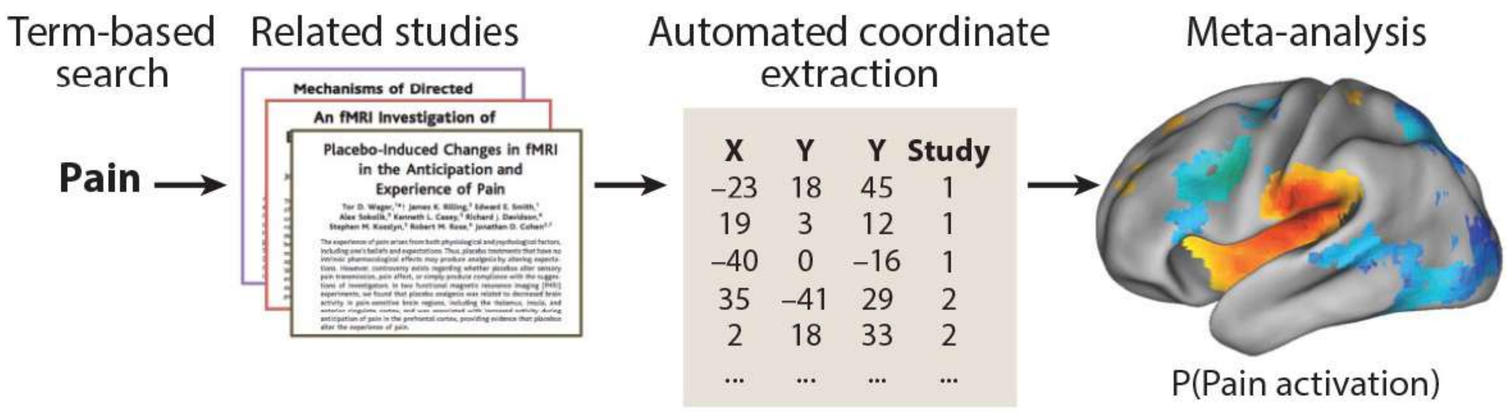
Why modeling cognition? We can look into brains!

- Does mind = brain?
- Definition of brain mapping:

"the creation of a visual representation of the brain in which different cognitive functions are assigned to different brain regions."
(American Psychological Association)

Why modeling cognition? We can look into brains!

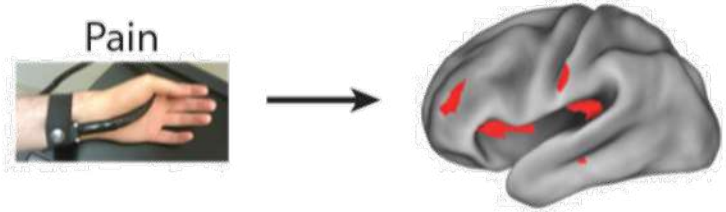
- Mapping psychological constructs on the brain is notoriously tricky



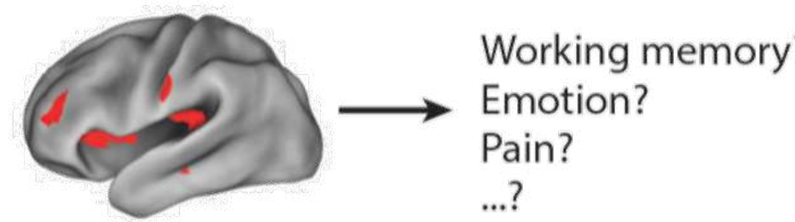
Yarkoni and Poldrack (2016) From brain maps to cognitive ontologies. *Annu. Rev. Psychol.*

Why modeling cognition?
We can look into brains!

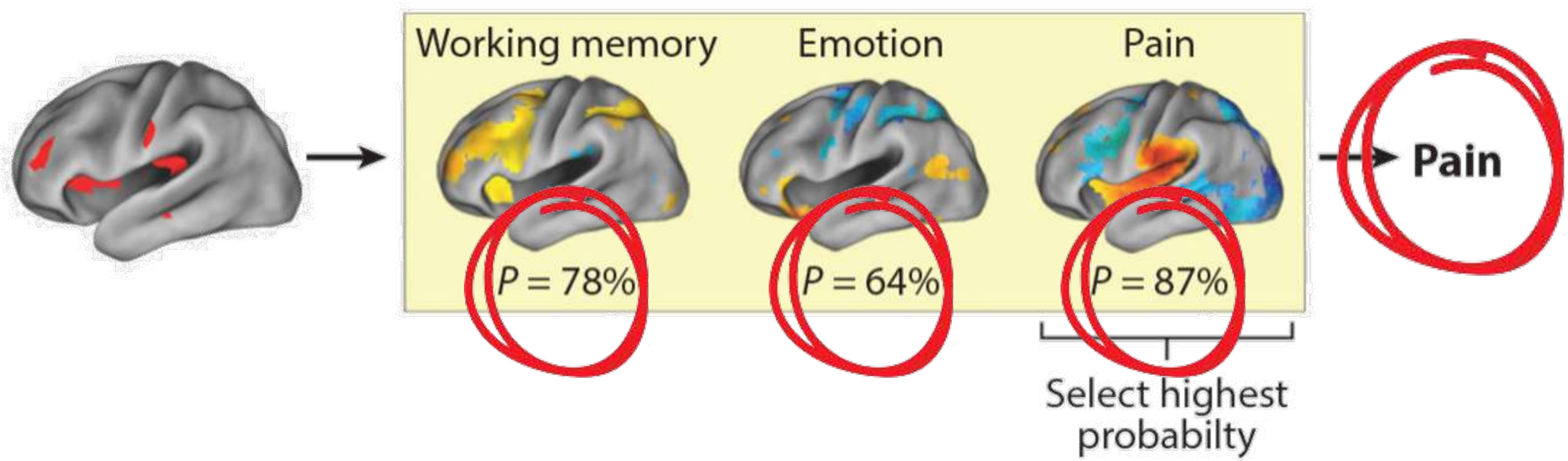
forward inference



reverse inference



classification (machine learning)



Why modeling cognition?

~~We can look into brains!~~

- Shimon Edelman: “the mind (as cognitive system) is best defined not in terms of its physical substrate, but in terms of the relations that states of the system have to one another, and to the outside world.”

- Correspondences can be identified between physically dissimilar cognitive systems in terms of shared computations.

Why modeling cognition?
~~We can look into brains!~~

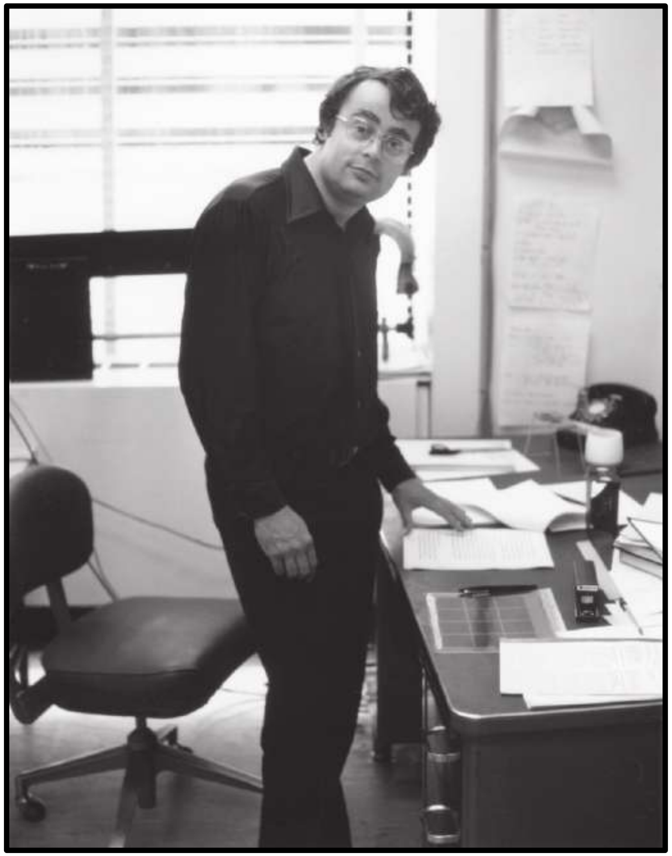
- Does mind = computer then?

not a physical one, obviously…

but a metaphorical one, arguably.

Why modeling cognition?

- David Marr's (1945-1980) three-level description of cognitive processes:
  - ✓ computation (upper level)
  - ✓ algorithm (middle level)
  - ✓ implementation (lower level)

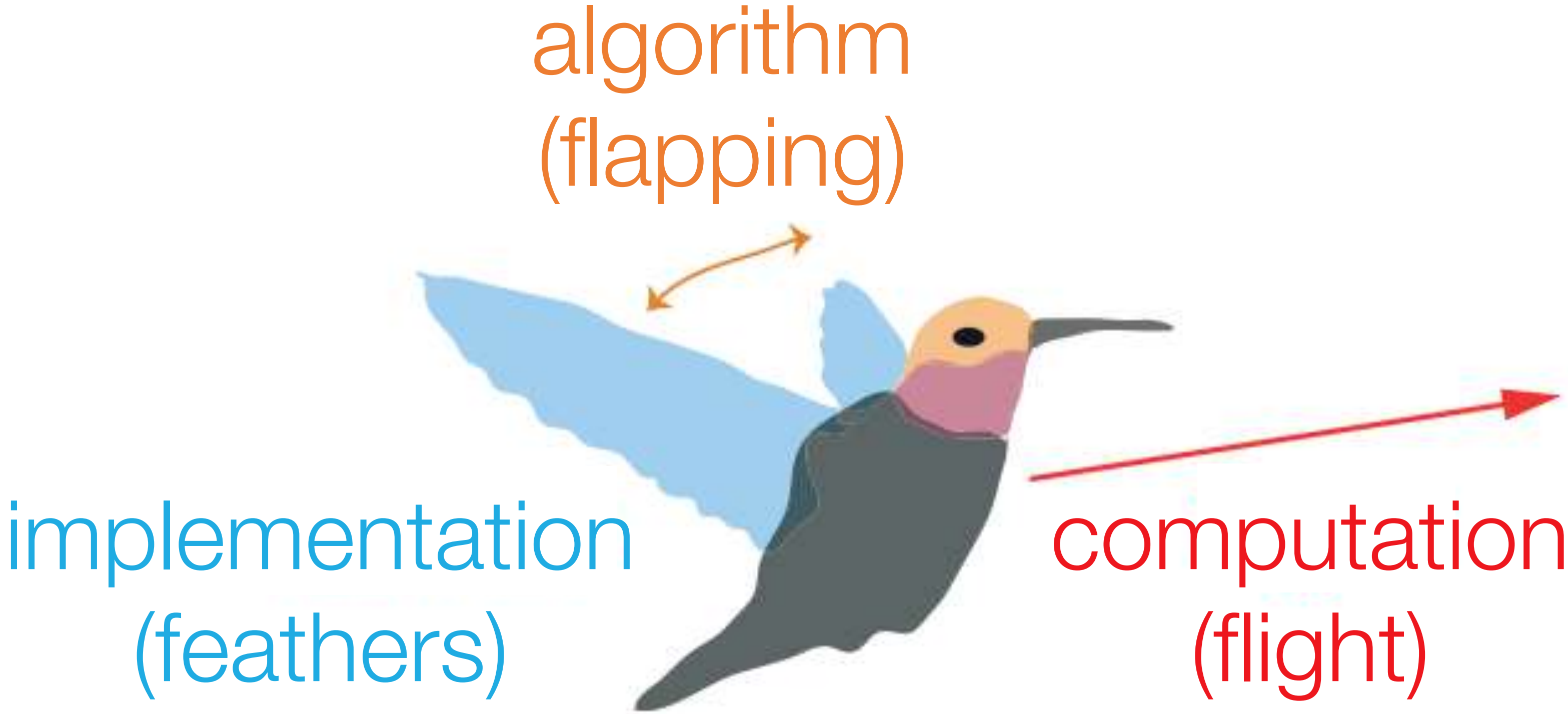


Why modeling cognition?

- Computation level:
  what does the system do?
- Algorithm level: (target of cognitive modeling)
  how does the system do it?
- Implementation level:
  what is the physical substrate of the system?

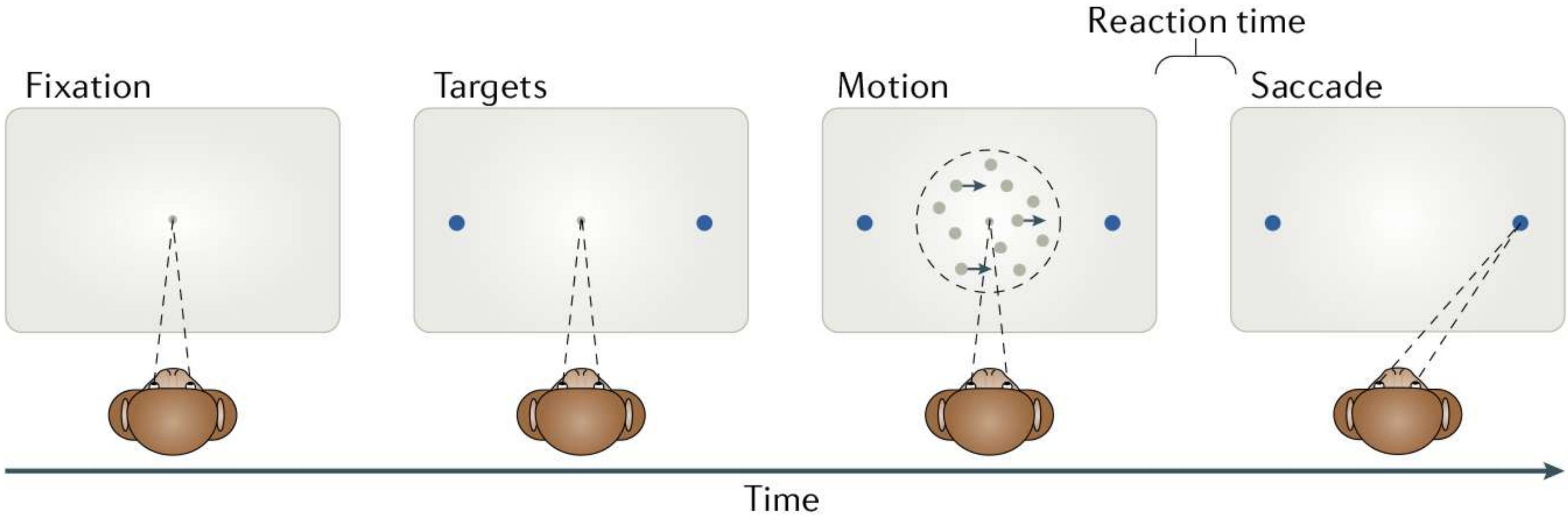
Marr and Poggio (1976) From understanding computation to understanding neural circuitry. *AI Memo Arch. from MIT AI Lab. (1959-2004)*

Why modeling cognition?



Why modeling cognition?

random-dot motion discrimination task

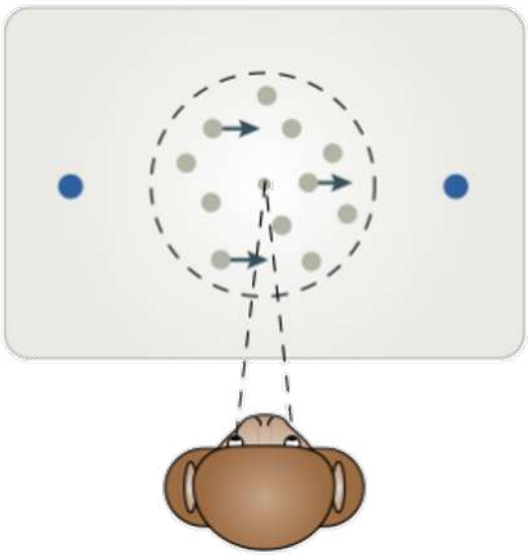


motion coherence
(from 0 to 100%)

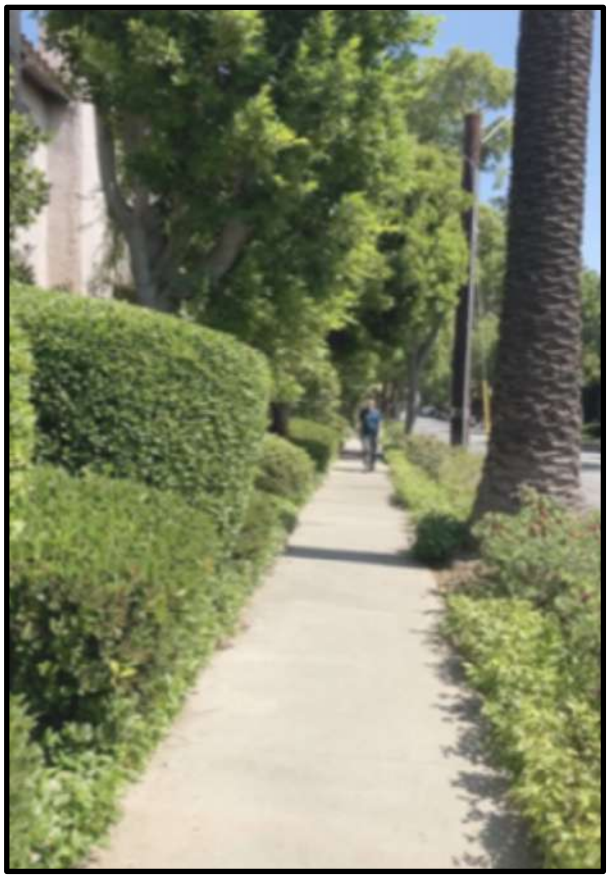
Why modeling cognition?

- Computation: Bayesian inference

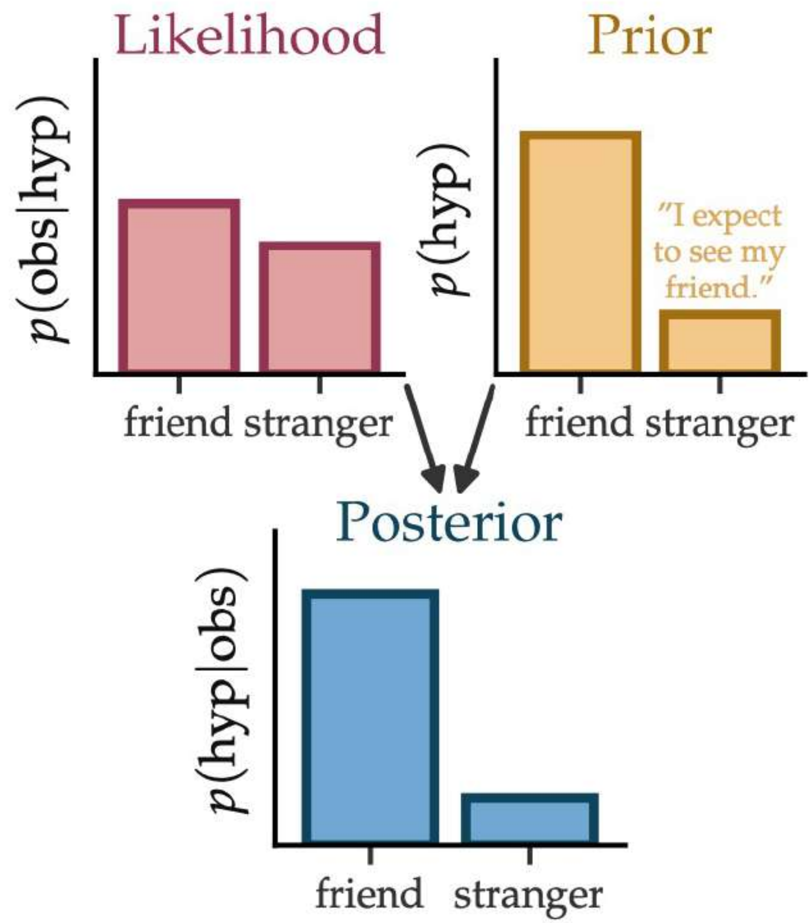
$$\underbrace{p(\text{hyp}|\text{obs})}_{\text{Posterior}} \propto \underbrace{p(\text{obs}|\text{hyp})}_{\text{Likelihood}} \cdot \underbrace{p(\text{hyp})}_{\text{Prior}}$$



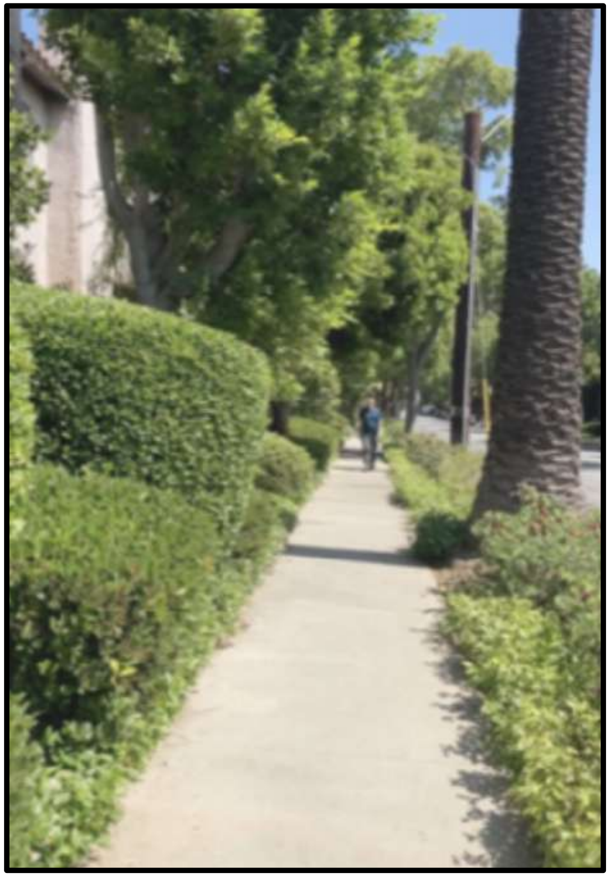
Why modeling cognition?



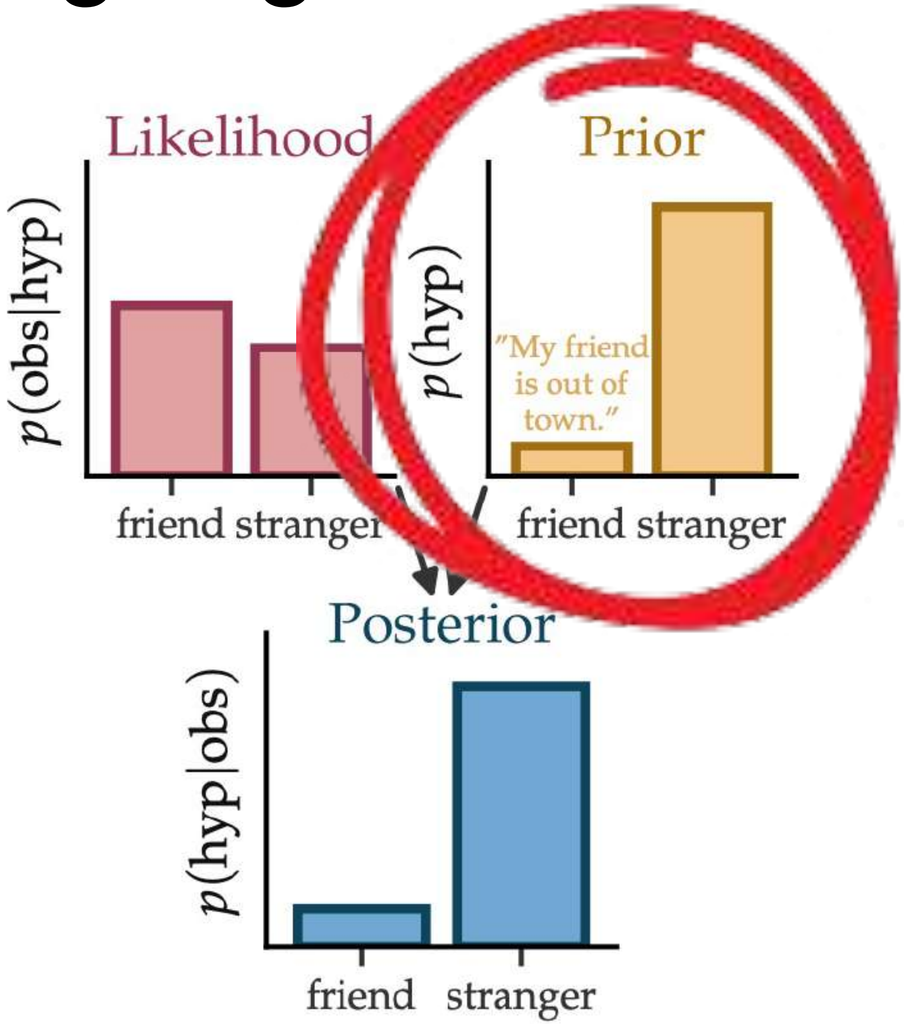
p(this is my friend)



Why modeling cognition?

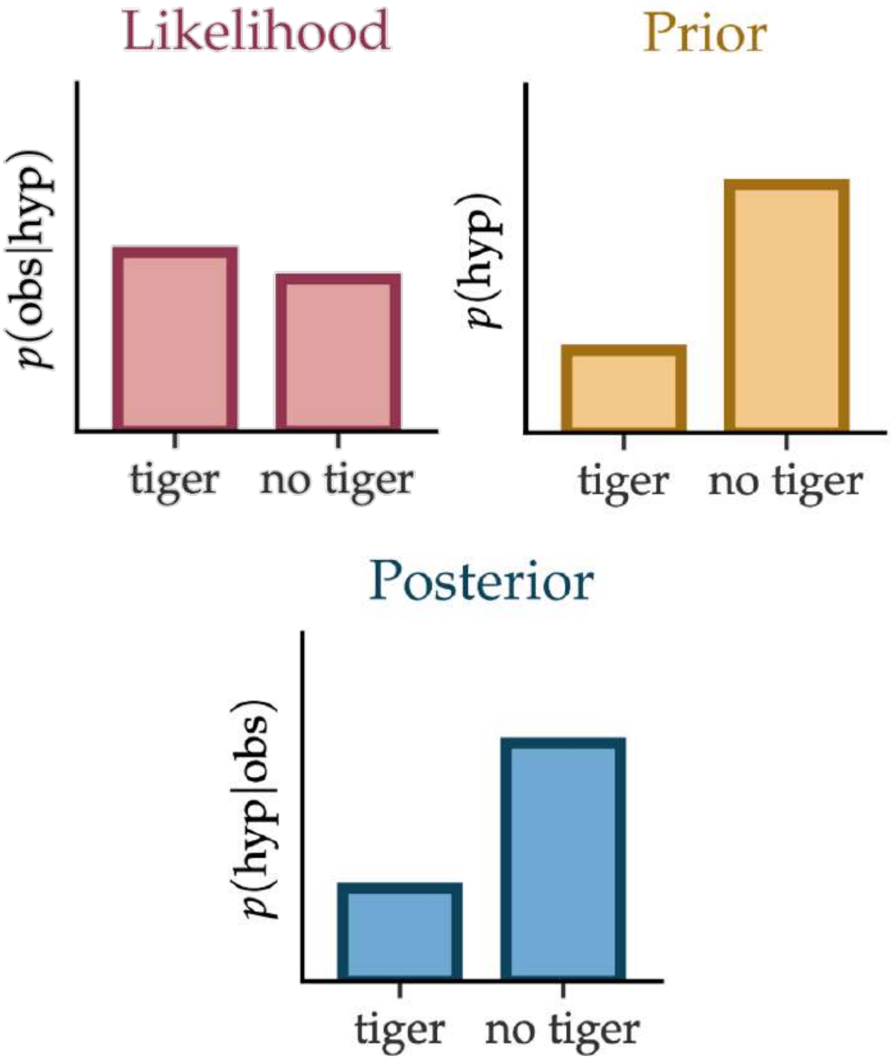


p(this is my friend)



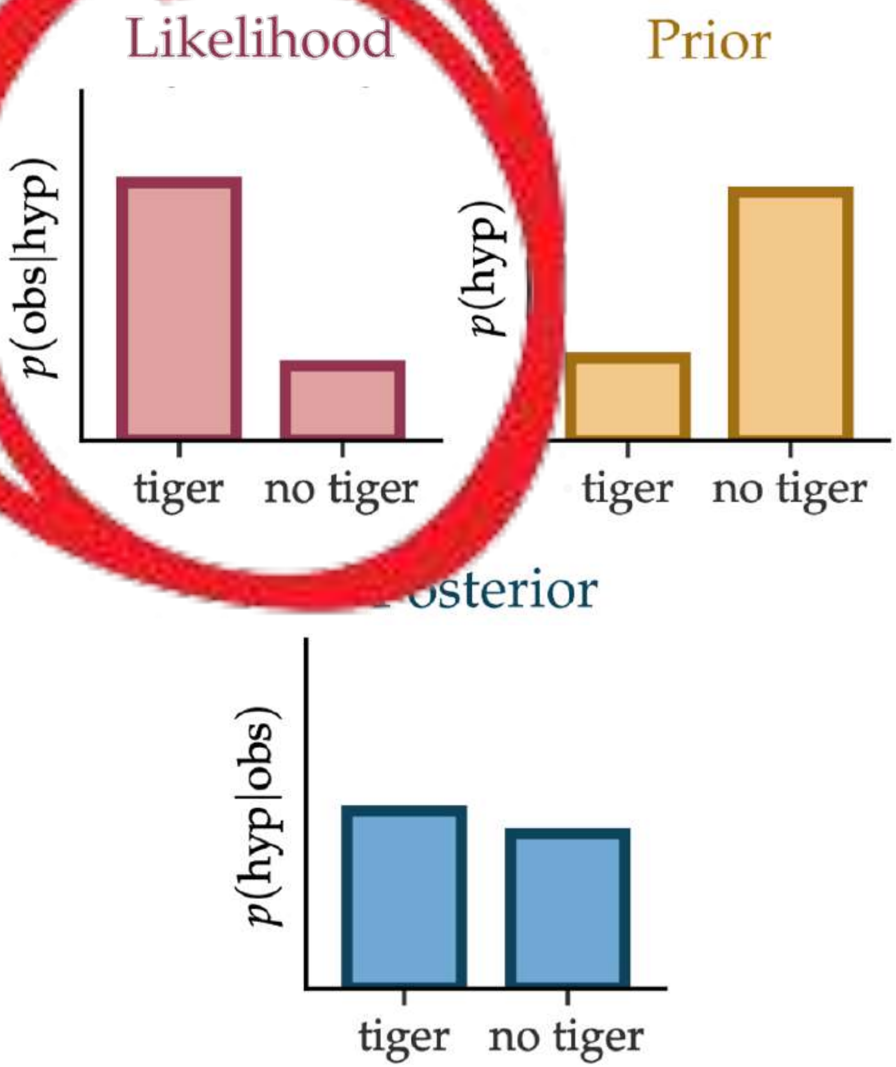
Why modeling cognition?

p(tiger in the bush)

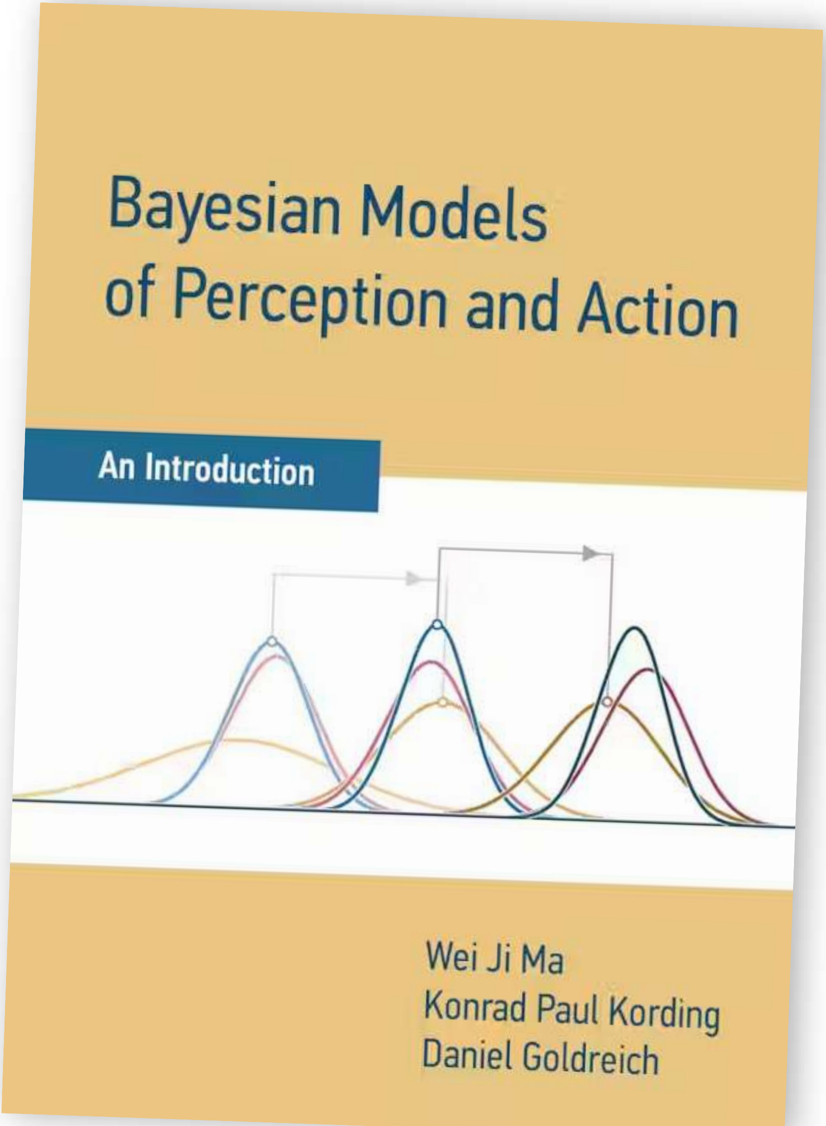


Why modeling cognition?

p(tiger in the bush)



Why modeling cognition?



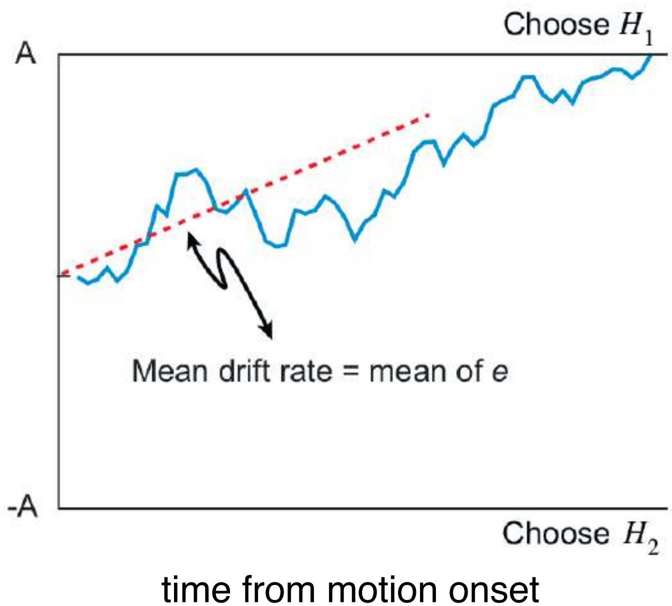
computation
(inference)

algorithm
(Bayes rule)

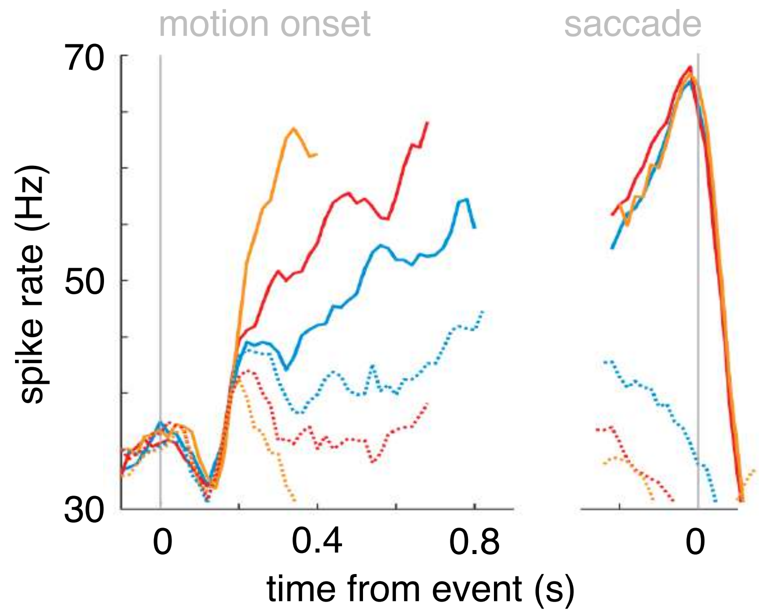
implementation
(population code)

Why modeling cognition?

Algorithm:
drift-diffusion model



Implementation:
LIP spike rate



Different flavors of models

- Model A:
  accuracy ~ 1 + motion coherence + RT
  (linear model)

- Model B:
  motion evidence accumulation-to-threshold
  (drift-diffusion model)

Different flavors of models

- Model A: descriptive
  accuracy ~ 1 + motion coherence + RT
  (linear model)

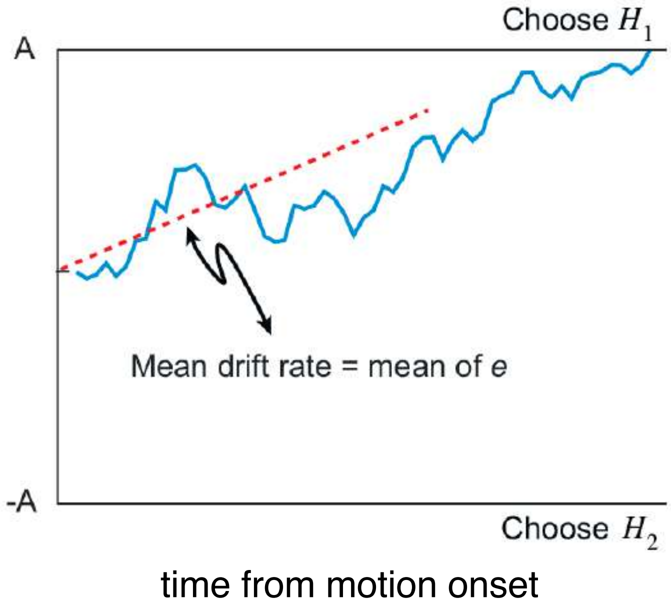
- Model B:
  motion evidence accumulation-to-threshold
  (drift-diffusion model)

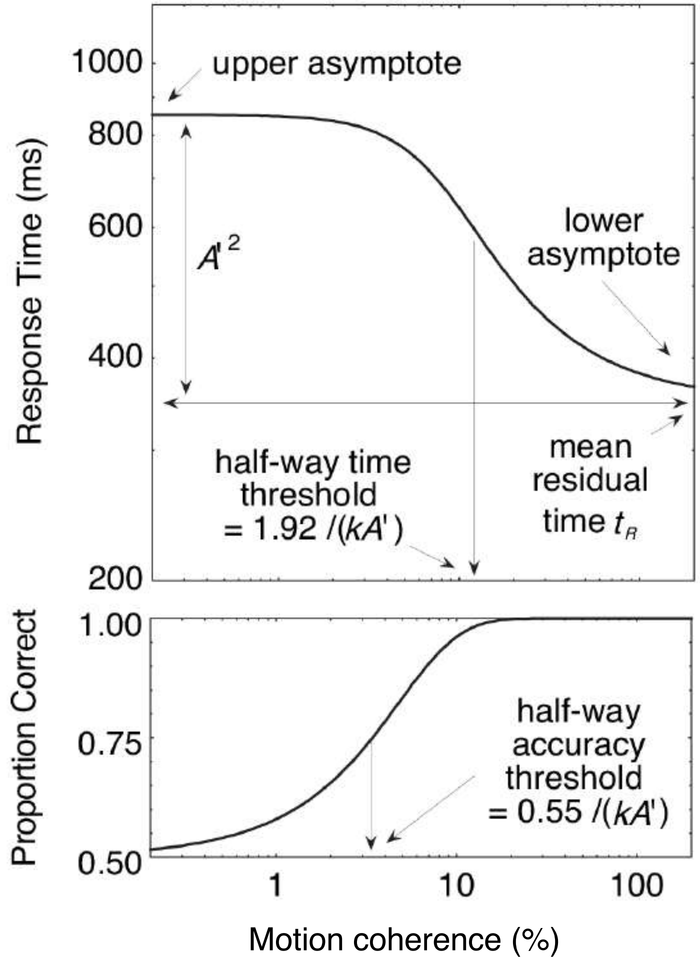
Different flavors of models

- Model A: descriptive
  accuracy ~ 1 + motion coherence + RT
  (linear model)
- Model B: generative
  motion evidence accumulation-to-threshold
  (drift-diffusion model)
- Cognitive modeling aims at building and testing generative models of behavior.

Different flavors of models

Model B:
drift-diffusion





Different flavors of models

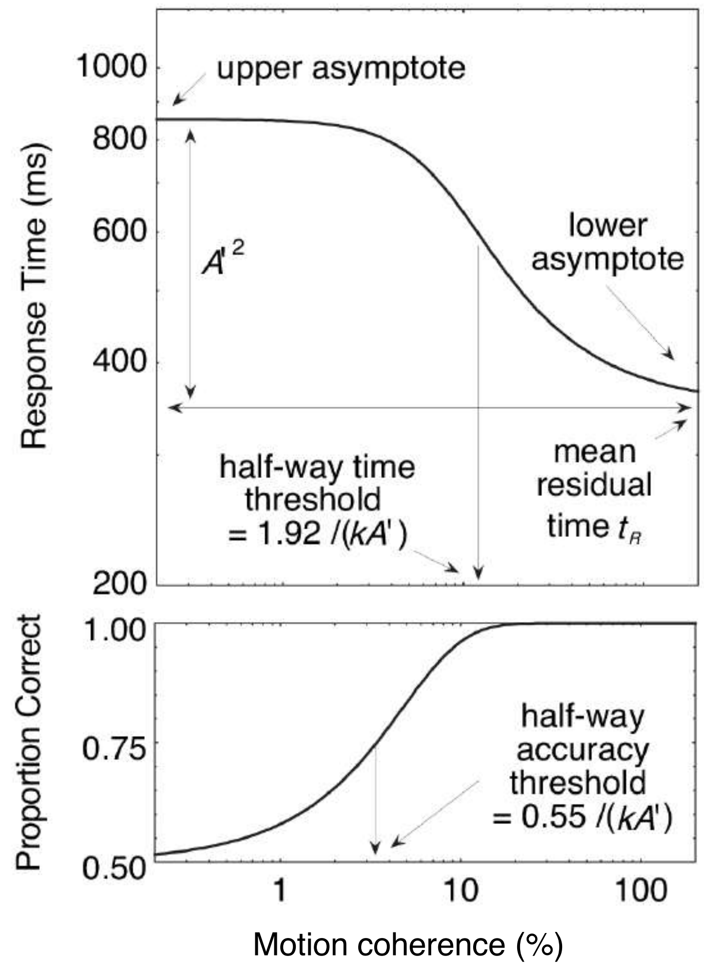
Model B:

chronometry

$$t_T(x) = \frac{A'}{kx}\tanh(A'kx) + t_R$$

accuracy

$$P_C(x) = \frac{1}{1+e^{-2A'k|x|}}$$



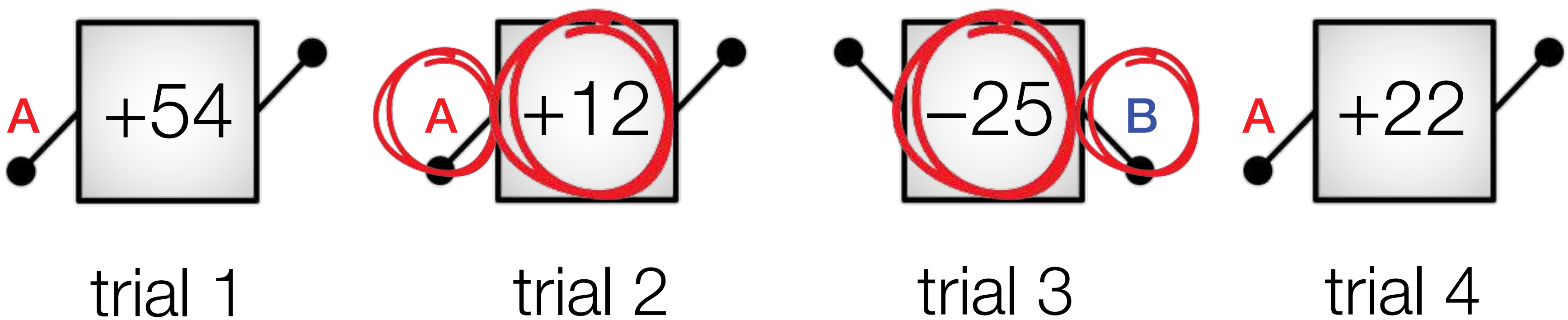
Class evaluation

- Create 5 groups of 3 students each
- In-depth analysis of behavioral data collected in a slot machine game (two-armed bandit)
- Use data mining <u>and</u> modeling approaches seen during lectures <u>and</u> practical sessions
- <u>Objective:</u> identify the latent cognitive strategy that drives behavior (<u>different</u> for each group)
- Group presentation (15 min/group) on Friday

Class evaluation

Class evaluation

two-armed bandit task



Paper to read

eLIFE
elifesciences.org

REVIEW ARTICLE

Ten simple rules for the computational modeling of behavioral data

Robert C Wilson[1,2†*], **Anne GE Collins**[3,4†*]

[1]Department of Psychology, University of Arizona, Tucson, United States; [2]Cognitive Science Program, University of Arizona, Tucson, United States; [3]Department of Psychology, University of California, Berkeley, Berkeley, United States; [4]Helen Wills Neuroscience Institute, University of California, Berkeley, Berkeley, United States

Abstract Computational modeling of behavior has revolutionized psychology and neuroscience. By fitting models to experimental data we can probe the algorithms underlying behavior, find neural correlates of computational variables and better understand the effects of drugs, illness and interventions. But with great power comes great responsibility. Here, we offer ten simple rules to ensure that computational modeling is used with care and yields meaningful insights. In particular, we present a beginner-friendly, pragmatic and details-oriented introduction on how to relate models to data. What, exactly, can a model tell us about the mind? To answer this, we apply our rules to the simplest modeling techniques most accessible to beginning modelers and illustrate them with examples and code available online. However, most rules apply to more advanced

Coming next

- Practical session: today, 2.00pm, same room

- Guidelines for cognitive modeling:
  Wilson and Collins (2019) Ten simple rules for the computational modeling of behavioral data. *eLife*
  https://doi.org/10.7554/eLife.49547 (open-access)

- Contact:

Valentin Wyart valentin.wyart@ens.psl.eu
Lucas Benjamin lucas.benjamin78@gmail.com

Lab. de Neurosciences Cognitives et Computationnelles (LNC^2)
Institut National de la Santé et de la Recherche Médicale
Ecole Normale Supérieure, Université PSL