

Psych 253

Advanced Statistical Modeling

What is Data? What's a Model?

Daniel Yamins

Departments of Psychology and Computer Science
Wu Tsai Neurosciences Institute
Stanford University

Russ Poldrack

Department of Psychology
Stanford University

Main Goals of Class

Learn:

How and why to build models of behavioral & neural data

Main Goals of Class

Learn:

How and why to build models of behavioral & neural data

How to evaluate and compare such models

Main Goals of Class

Learn:

How and why to build models of behavioral & neural data

There is a unified framework for capturing most models ... not just a series of arbitrary “named methods”

How to evaluate and compare such models

Main Goals of Class

Learn:

How and why to build models of behavioral & neural data

There is a unified framework for capturing most models ... not just a series of arbitrary “named methods”

Motivate & construct a series of especially useful simple models in this framework

How to evaluate and compare such models

Main Goals of Class

Learn:

How and why to build models of behavioral & neural data

There is a unified framework for capturing most models ... not just a series of arbitrary “named methods”

Motivate & construct a series of especially useful simple models in this framework

How to use several key general modeling packages in a systematic fashion

How to evaluate and compare such models

Main Goals of Class

Learn:

How and why to build models of behavioral & neural data

There is a unified framework for capturing most models ... not just a series of arbitrary “named methods”

Motivate & construct a series of especially useful simple models in this framework

How to use several key general modeling packages in a systematic fashion

How to evaluate and compare such models

There is a unified framework for training/testing/comparing models ... not just a series of arbitrary statistical tests.

Main Goals of Class

Learn:

How and why to build models of behavioral & neural data

There is a unified framework for capturing most models ... not just a series of arbitrary “named methods”

Motivate & construct a series of especially useful simple models in this framework

How to use several key general modeling packages in a systematic fashion

How to evaluate and compare such models

There is a unified framework for training/testing/comparing models ... not just a series of arbitrary statistical tests.

Simple, flexible, robust tools for performing model evaluation & comparison

Main Goals of Class

Learn:

How and why to build models of behavioral & neural data

There is a unified framework for capturing most models ... not just a series of arbitrary “named methods”

Motivate & construct a series of especially useful simple models in this framework

How to use several key general modeling packages in a systematic fashion

How to evaluate and compare such models

There is a unified framework for training/testing/comparing models ... not just a series of arbitrary statistical tests.

Simple, flexible, robust tools for performing model evaluation & comparison

How to fairly compute **data** reliability and quantify experimental limitations

What is Data?

What is a model (of behavioral or neural data)?

What is Data?

What is a model (of behavioral or neural data)?



What is (behavioral or neural) data?

What is Data?

Def'n: Data is a set of linked tensors.

What is Data?

Def'n: Data is a set of linked tensors.

Vectors = list of numbers

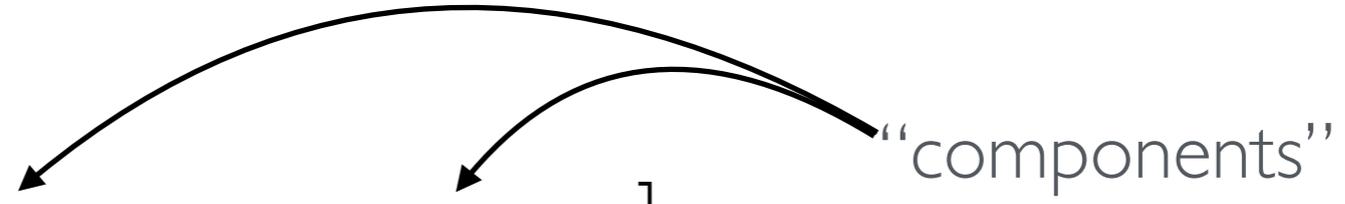
$$\mathbf{v} = [0.4, 0.2, -1.144, \dots, 9.5]$$

What is Data?

Def'n: Data is a set of linked tensors.

Vectors = list of numbers

$$\mathbf{v} = [0.4, 0.2, -1.144, \dots, 9.5]$$

$$\mathbf{v} = [v_0, v_1, v_2, \dots, v_{N-1}]$$


“components”

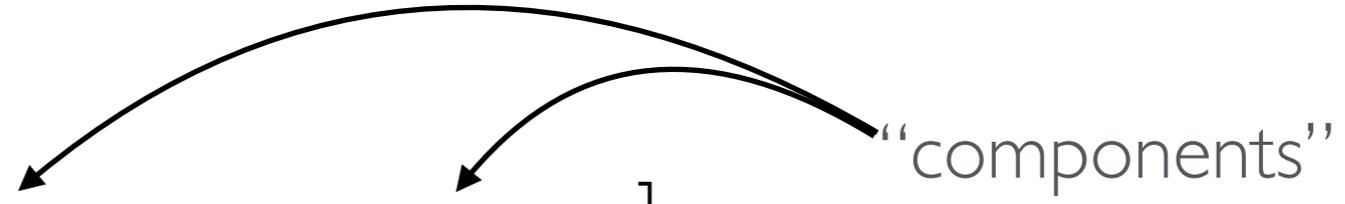
The diagram consists of two curved arrows pointing downwards from the numerical vector above to the symbolic vector below. The first arrow starts under the value '0.4' and points to the symbol 'v₁'. The second arrow starts under the value '9.5' and points to the symbol 'v_{N-1}'. The text 'components' is positioned to the right of the second arrow, with a line pointing to the arrow's tip.

What is Data?

Def'n: Data is a set of linked tensors.

Vectors = list of numbers

$$\mathbf{v} = [0.4, 0.2, -1.144, \dots, 9.5]$$

$$\mathbf{v} = [v_0, v_1, v_2, \dots, v_{N-1}]$$


The diagram consists of two curved arrows pointing from the text "components" on the right to the terms v_1 and v_{N-1} in the vector notation below. A longer arrow also points from the text "components" to the first term v_0 in the same notation.

N = "dimension" of the data = length of the list

What is Data?

Def'n: Data is a set of linked tensors.

Matrices = 2-dimensional array of numbers

$$M = \begin{bmatrix} 1.1 & -0.3 & -0.22 & 0.64 & 1.9 \\ 9.1 & -0.91 & 0.31 & 0.441 & 10.3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 3.2 & -0.131 & 1.5 & \dots & 0.3 \end{bmatrix}$$

What is Data?

Def'n: Data is a set of linked tensors.

Matrices = 2-dimensional array of numbers

$$M = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & x_{m3} & \dots & x_{mn} \end{bmatrix}$$

(m, n) = “shape” of the data.

What is Data?

Def'n: Data is a set of linked tensors.

Vector

't'
'e'
'n'
's'
'o'
'r'

1-Tensor

Matrix

3	1	4	1
5	9	2	6
5	3	5	8
9	7	9	3
2	3	8	4
6	2	6	4

2-Tensor

Data Cube

...

3-Tensor

$$T[i_1, i_2, \dots, i_D] = \text{some number}$$

D = dimension of the tensor

(n_1, \dots, n_D) = shape

What is Data?

Def'n: Data is a set of linked tensors.

Color Image = typically, a 3 Tensor



two bears.jpg

What is Data?

Def'n: Data is a set of linked tensors.

Color Image = typically, a 3 Tensor



two bears.jpg

[IPYNB:“Loading Images as Python Arrays”]

What is Data?

Def'n: Data is a set of linked tensors.

Color Image = typically, a 3 Tensor



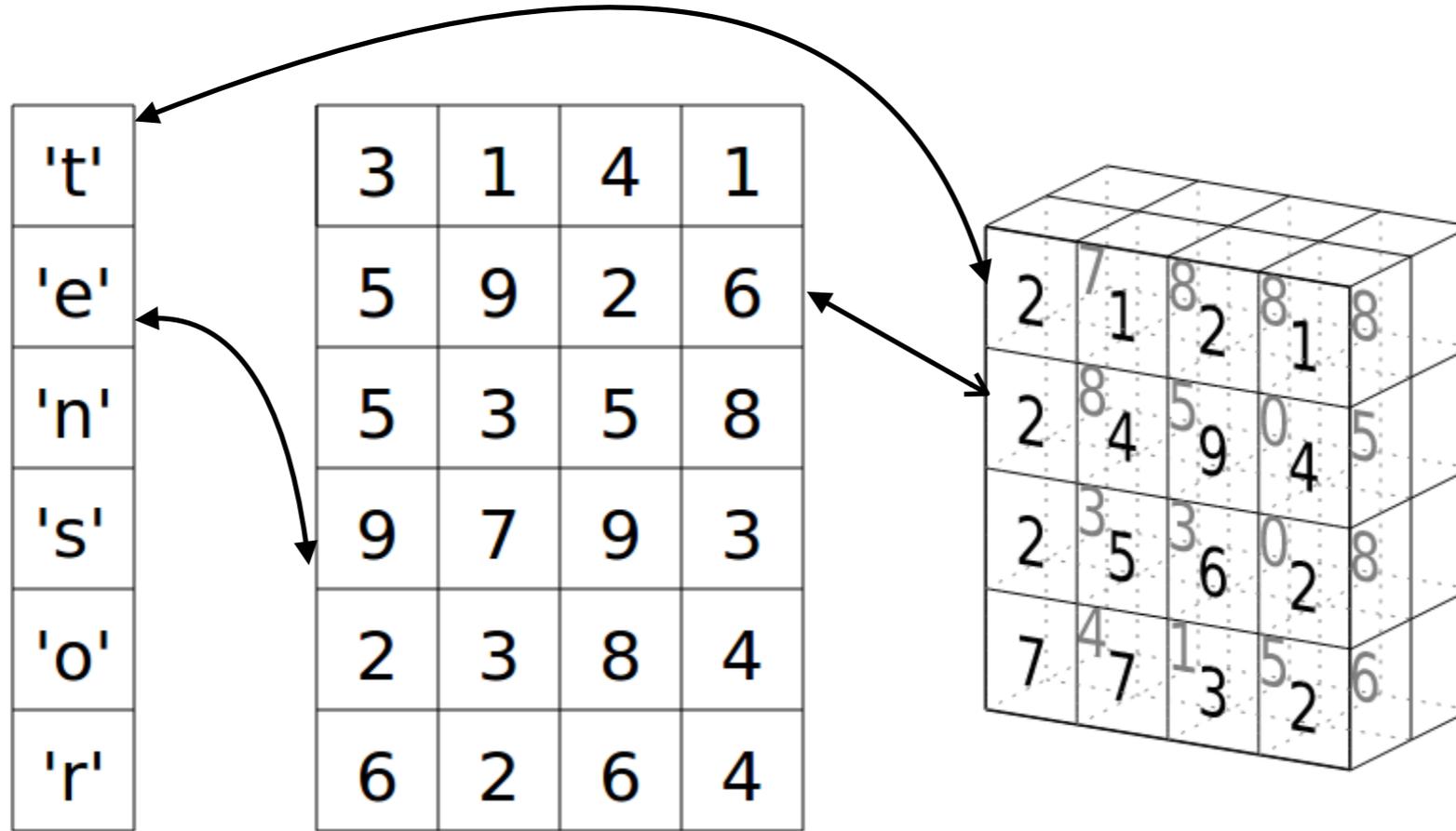
two bears.jpg

[IPYNB:“Loading Images as Python Arrays”]

3-Tensor with shape
(194, 260, 3)
rows cols color channels

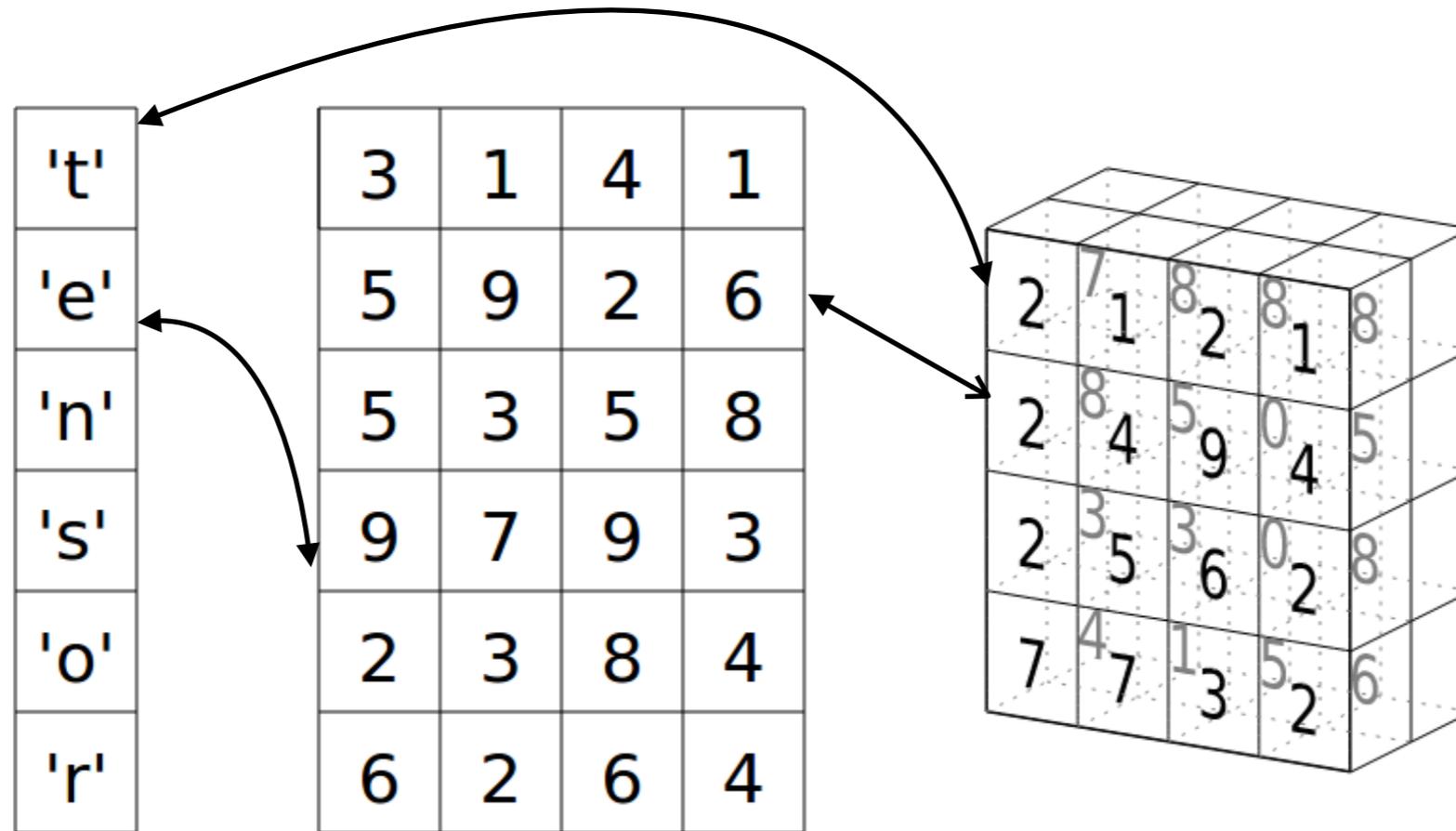
What is Data?

Def'n: Data is a set of **linked** tensors.



What is Data?

Def'n: Data is a set of **linked** tensors.



Let's explain what this means in terms of some real examples.

Example I: Self-Regulation Ontology

The Self-Regulation Ontology Project



Ian Eisenberg



Russ Poldrack

Example I: Self-Regulation Ontology

Self-Regulation: “ability to modulate behavior in service of long-term goals.”

Example I: Self-Regulation Ontology

Self-Regulation: “ability to modulate behavior in service of long-term goals.”

Behaviour Research and Therapy 101 (2018) 46–57



Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Applying novel technologies and methods to inform the ontology of self-regulation

Ian W. Eisenberg^{a,*}, Patrick G. Bissett^a, Jessica R. Canning^d, Jesse Dallery^f, A. Zeynep Enkavi^a, Susan Whitfield-Gabrieli^j, Oscar Gonzalez^d, Alan I. Green^b, Mary Ann Greene^c, Michaela Kiernan^e, Sunny Jung Kim^{b,c}, Jamie Li^a, Michael R. Lowe^k, Gina L. Mazza^d, Stephen A. Metcalf^c, Lisa Onkenⁱ, Sadev S. Parikh^a, Ellen Peters^g, Judith J. Prochaska^e, Emily A. Scherer^b, Luke E. Stoeckel^h, Matthew J. Valente^d, Jialing Wu^c, Haiyi Xie^c, David P. MacKinnon^d, Lisa A. Marsch^{b,c}, Russell A. Poldrack^a

^a Department of Psychology, Stanford University, Stanford, CA 94305, USA

^b Department of Psychiatry, Geisel School of Medicine at Dartmouth, Dartmouth College, Lebanon, NH 03766, USA

^c Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Dartmouth College, Lebanon, NH 03766, USA

^d Department of Psychology, Arizona State University, Tempe, AZ 85281, USA

^e Department of Medicine, Stanford University, Stanford, CA 94305, USA

^f Department of Psychology, University of Florida, Gainesville, FL 32611, USA

^g Department of Psychology, The Ohio State University, Columbus, OH 43206, USA

^h National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

ⁱ National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA

^j Brain and Cognitive Sciences Department, The McGovern Institute for Brain Research, MIT, Cambridge, MA 02139, USA

^k Department of Psychology, Drexel University, Philadelphia, PA, USA

Example I: Self-Regulation Ontology

The Self-Regulation Ontology Project:

Measure three things:

- 1) People's performance on cognitive laboratory tasks thought to be related to self-regulatory behavior

Example I: Self-Regulation Ontology

The Self-Regulation Ontology Project:

Measure three things:

- 1) People's performance on cognitive laboratory tasks thought to be related to self-regulatory behavior
- 2) People's self-reported responses to surveys about their lives

Example I: Self-Regulation Ontology

The Self-Regulation Ontology Project:

Measure three things:

- 1) People's performance on cognitive laboratory tasks thought to be related to self-regulatory behavior
- 2) People's self-reported responses to surveys about their lives
- 3) Actual outcomes for a bunch of important health variables

Example I: Self-Regulation Ontology

The Self-Regulation Ontology Project:

Measure three things:

- 1) People's performance on cognitive laboratory tasks thought to be related to self-regulatory behavior
- 2) People's self-reported responses to surveys about their lives
- 3) Actual outcomes for a bunch of important health variables

Q: What aspects of 1) and 2) predict which aspects of 3)?

Table 1
Cognitive tasks.

Task	Individual Difference Variables	References	Go-NoGo	D-prime Bias Score	
Adaptive N-Back	DDM Parameters Average load	(Harvey et al., 2005; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008)	Hierarchical Learning Task Holt & Laury	Attention to Hierarchy Percent Patient	(Badre, Kayser, & D'Esposito, 2010) (Holt & Laury, 2002)
Angling Risk Task	Two Conditions (Keep, Release): Adjusted Clicks Loss Percent Score	(Pleskac, 2008)	Information Sampling Task	Two conditions (Decreasing Win, Fixed Win): Probability Correct at choice Motivation	(Clark, Robbins, Ersche, & Sahakian, 2006)
Attention Network Task	DDM Parameters Alerting Effect Orient Effect Conflict Effect	(Fan, McCandliss, Fossella, Flombaum, & Posner, 2005)	Keep Track Task	Score	(A. Miyake et al., 2000; Yntema, 1963)
Bickel Titrator	Discount Rate for three payout magnitudes	(Koffarnus & Bickel, 2014)	Kirby	Discount Rate for three payout magnitudes	(Kirby & Maraković, 1996)
Choice Reaction Time Cognitive Reflection Task	DDM Parameters Correct Proportion Intuitive Proportion	(Primi, Morsanyi, Chiesi, Donati, & Hamilton, 2016; Toplak, West, & Stanovich, 2014)	Local-Global	DDM Parameters Switch Cost Conflict Effect	(A. Miyake et al., 2000; Yntema, 1963)
Columbia Card Task Cold/Hot	Average # of cards chosen Gain Sensitivity Loss Sensitivity # Loss Cards Sensitivity Level of Information Use	(Figner, Mackinlay, Wilkening, & Weber, 2009)	Motor Selective Stop Signal	DDM Parameters SSRT Reactive Control Selective Proactive Control	(Aron, Behrens, Smith, Frank, & Poldrack, 2007)
Dietary Decision Task	Health Sensitivity Taste Sensitivity	(Hare, Camerer, & Rangel, 2009)	Probabilistic Selection Task Psychological Refractory Period	Positive Learning Bias Value Sensitivity Slope of PRP function	(Frank, Seeberger, & O'Reilly, 2004) (Pashler, 1994)
Digit Span	Forward Span Reverse Span	(Woods et al., 2011)	Raven's Matrices Recent Probes	Score DDM Parameters Proactive Interference	(Raven & Raven, 2003) (Nee et al., 2007)
Directed Forgetting	DDM Parameters Proactive Interference	(Nee, Jonides, & Berman, 2007)	Shape Matching Task	DDM Parameters Stimulus Interference	(Stahl et al., 2014)
Discount Titrator	Percent Patient	(Figner et al., 2010)			
Dot Pattern Expectancy	DDM Parameters AY-BY BX-BY D-prime Bias	(Otto, Skatova, Madlon-Kay, & Daw, 2013)			

Table 1

Cognitive tasks.

Task	Individual Difference Variables	References
Adaptive N-Back	DDM Parameters Average load	(Harvey et al., 2005; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008)
Angling Risk Task	Two Conditions (Keep, Release): Adjusted Clicks Loss Percent Score	(Pleskac, 2008)
Attention Network Task	DDM Parameters Alerting Effect Orient Effect Conflict Effect	(Fan, McCandliss, Fossella, Flombaum, & Posner, 2005)
Bickel Titrator	Discount Rate for three payout magnitudes	(Koffarnus & Bickel, 2014)
Choice Reaction Time Cognitive Reflection Task	DDM Parameters Correct Proportion Intuitive Proportion	(Primi, Morsanyi, Chiesi, Donati, & Hamilton, 2016; Toplak, West, & Stanovich, 2014)
Columbia Card Task Cold/Hot	Average # of cards chosen Gain Sensitivity Loss Sensitivity # Loss Cards Sensitivity Level of Information Use	(Figner, Mackinlay, Wilkening, & Weber, 2009)
Dietary Decision Task	Health Sensitivity Taste Sensitivity	(Hare, Camerer, & Rangel, 2009)
Digit Span	Forward Span Reverse Span	(Woods et al., 2011)
Directed Forgetting	DDM Parameters Proactive Interference	(Nee, Jonides, & Berman, 2007)
Discount Titrator	Percent Patient	(Figner et al., 2010)
Dot Pattern Expectancy	DDM Parameters AY-BY BX-BY D-prime Bias	(Otto, Skatova, Madlon-Kay, & Daw, 2013)

Table 1 (continued)

Task	Individual Difference Variables	References
Shift Task	Accuracy Learning Rate	(Wilson & Niv, 2012)
Simon Task	DDM Parameters Simon Effect	(Lu & Proctor, 1995)
Simple Reaction Time	Average Reaction Time	
Spatial Span	Forward Span Reverse Span	(Woods et al., 2011)
Stimulus Selective Stop Signal	DDM Parameters SSRT	(Bissett & Logan, 2014)
Stop Signal	DDM Parameters SSRT (low stop signal probability condition) SSRT (high stop signal probability condition) Proactive SSRT speeding Proactive Slowing	(Bissett & Logan, 2011)
Stroop	DDM Parameters Stroop Effect	(A. Miyake et al., 2000; Yntema, 1963)
Three-By-Two	DDM Parameters Stimulus Switch Cost (100/900 CSI*) Task Switch Cost (100/900 CSI*) * Cue-Stimulus Interval	(Schneider & Logan, 2011)
Tower of London	Average Move Time # Extra Moves # Optimal Solutions Planning Time	(Shallice, 1982)
Two-Step Decision	Model-Based Index Model-Free Index Perseverance	(Daw, Gershman, Seymour, Dayan, & Dolan, 2011)
Writing Task	Positive Probability Negative Probability	

Go-NoGo	----- D-prime Bias Score
Hierarchical Learning Task	Attention to Hierarchy
Holt & Laury	Percent Patient
Information Sampling Task	Two conditions (Decreasing Win, Fixed Win): Probability Correct at choice Motivation
Keep Track Task	Score
Kirby	Discount Rate for three payout magnitudes
Local-Global	DDM Parameters Switch Cost Conflict Effect
Motor Selective Stop Signal	DDM Parameters SSRT Reactive Control Selective Proactive Control
Probabilistic Selection Task	Positive Learning Bias Value Sensitivity
Psychological Refractory Period	Slope of PRP function
Raven's Matrices	Score
Recent Probes	DDM Parameters Proactive Interference
Shape Matching Task	DDM Parameters Stimulus Interference

Triangulating a Cognitive Control Network Using Diffusion-Weighted Magnetic Resonance Imaging (MRI) and Functional MRI

Adam R. Aron, Tim E. Behrens, Steve Smith, Michael J. Frank, and Russell A. Poldrack

Journal of Neuroscience 4 April 2007, 27 (14) 3743-3752; DOI: <https://doi.org/10.1523/JNEUROSCI.0519-07.2007>

Abstract

The ability to stop motor responses depends critically on the right inferior frontal cortex (IFC) and also engages a midbrain region consistent with the subthalamic nucleus (STN). Here we used diffusion-weighted imaging (DWI) tractography to show that the IFC and the STN region are connected via a white matter tract, which could underlie a “hyperdirect” pathway for basal ganglia control. Using a novel method of “triangulation” analysis of tractography data, we also found that both the IFC and the STN region are connected with the presupplementary motor area (preSMA). We hypothesized that the preSMA could play a conflict detection/resolution role within a network between the preSMA, the IFC, and the STN region. A second experiment tested this idea with functional magnetic resonance imaging (fMRI) using a conditional stop-signal paradigm, enabling examination of behavioral and neural signatures of conflict-induced slowing. The preSMA, IFC, and STN region were significantly activated the greater the conflict-induced slowing. Activation corresponded strongly with spatial foci predicted by the DWI tract analysis, as well as with foci activated by complete response inhibition. The results illustrate how tractography can reveal connections that are verifiable with fMRI. The results also demonstrate a three-way functional–anatomical network in the right hemisphere that could either brake or completely stop responses.

Table 1 (continued)

Task	Individual Difference Variables	References
Shift Task	Accuracy Learning Rate	(Wilson & Niv, 2012)
Simon Task	DDM Parameters Simon Effect	(Lu & Proctor, 1995)
Simple Reaction Time	Average Reaction Time	
Spatial Span	Forward Span	(Woods et al., 2011)
Stimulus Selective Stop Signal		
Stop Signal		
Stroop		
Three-By-Two		
Tower of London	Average Move Time # Extra Moves # Optimal Solutions Planning Time	(Shallice, 1982)
Two-Step Decision	Model-Based Index Model-Free Index Perseverance	(Daw, Gershman, Seymour, Dayan, & Dolan, 2011)
Writing Task	Positive Probability Negative Probability	

ORIGINAL RESEARCH ARTICLE

Front. Hum. Neurosci., 24 January 2012 | <https://doi.org/10.3389/fnhum.2011.00189>

Inferring relevance in a changing world

Robert C. Wilson* and Yael Niv

Department of Psychology, Neuroscience Institute, Princeton University, Princeton, NJ, USA

STIMULUS SWITCH COST
(100/900 CSI*)

Task Switch Cost
(100/900 CSI*)

* Cue-Stimulus
Interval

Example 1: Self-Regulation Ontology

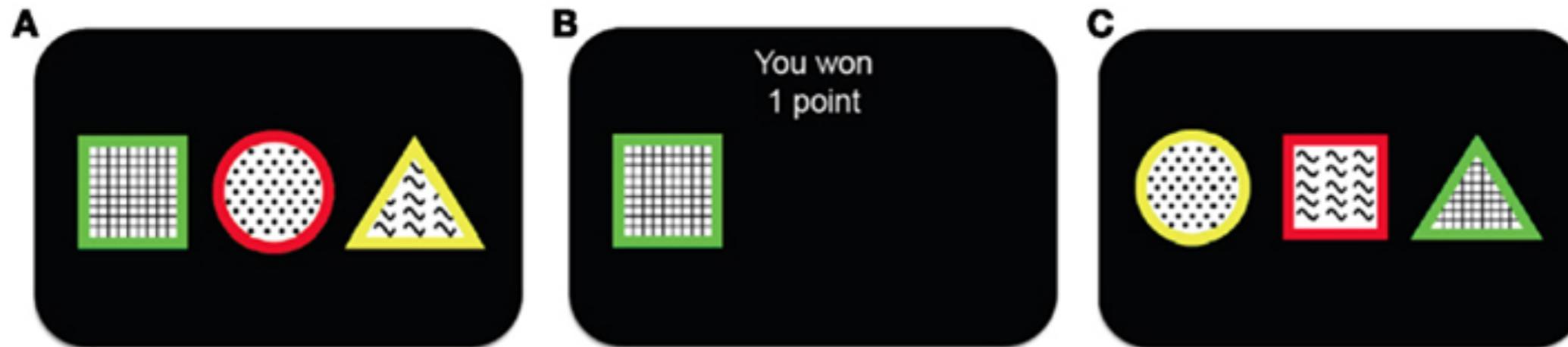


Figure 1. Schematic showing the outline of the task. (A) The subject is presented with three different stimuli. Each stimulus has a different feature along each one of the three feature dimensions (shape, color and texture). (B) The subject chooses one of the stimuli and receives binary reward feedback, either winning one (as in this case) or zero points. (C) After a short delay a new trial begins and the subject is presented with three new stimuli from which to choose.

Example I: Self-Regulation Ontology

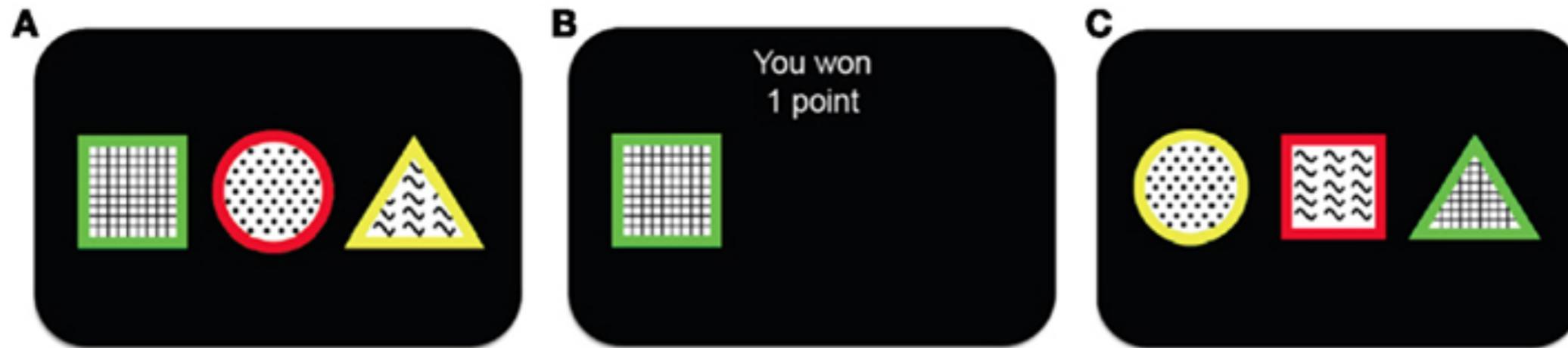


Figure 1. Schematic showing the outline of the task. (A) The subject is presented with three different stimuli. Each stimulus has a different feature along each one of the three feature dimensions (shape, color and texture). (B) The subject chooses one of the stimuli and receives binary reward feedback, either winning one (as in this case) or zero points. (C) After a short delay a new trial begins and the subject is presented with three new stimuli from which to choose.

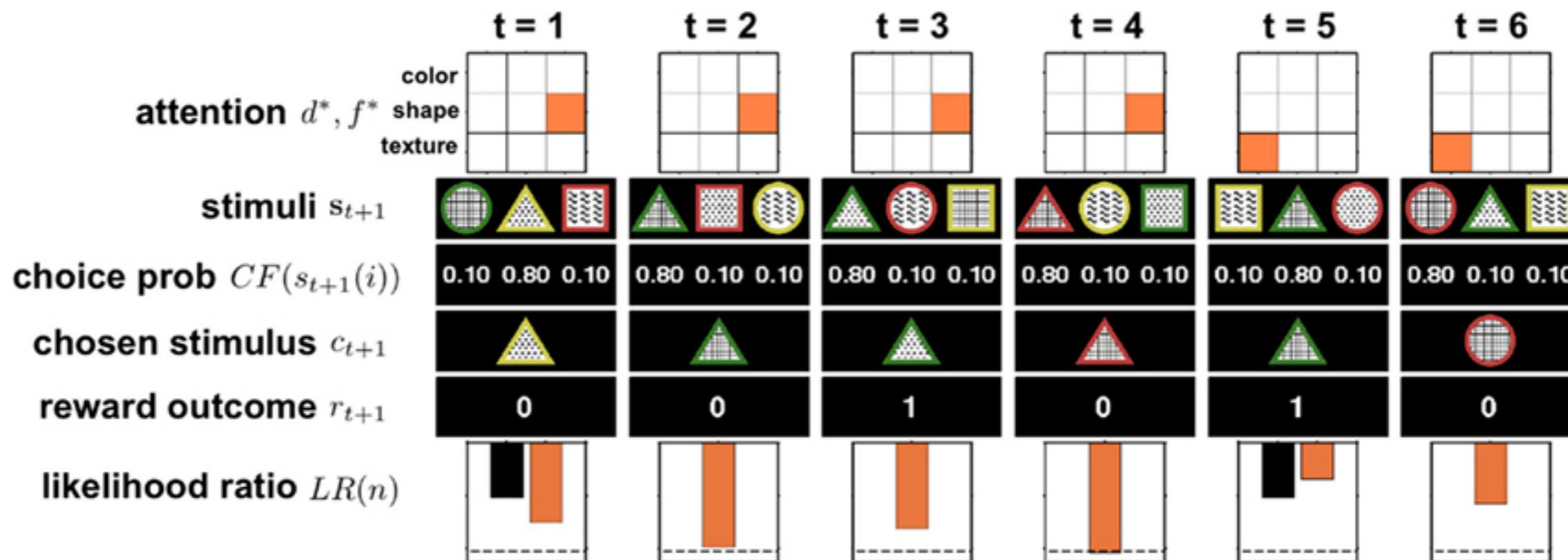


Figure 4: Illustration of six trials of the full selective attention model in action.

Example I: Self-Regulation Ontology

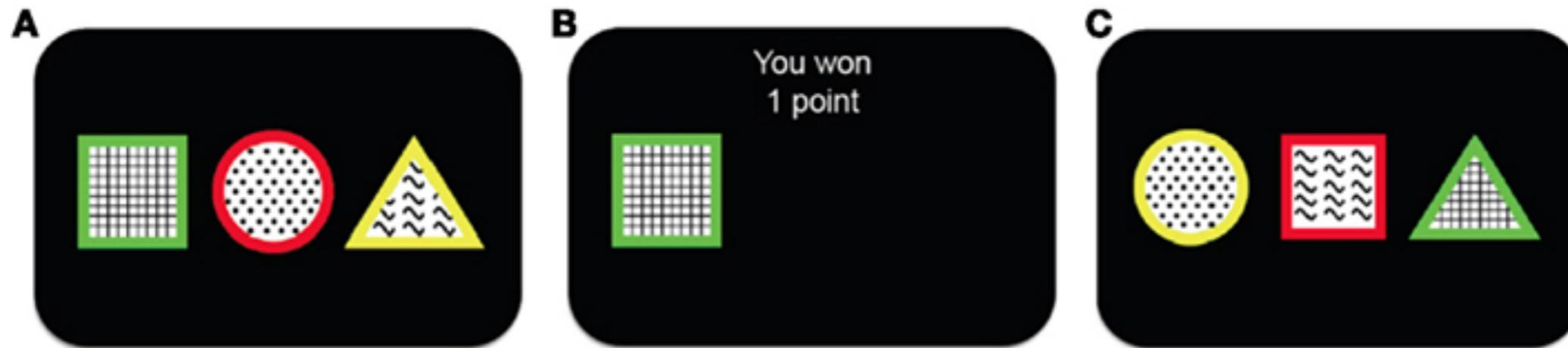


Figure 1. Schematic showing the outline of the task. (A) The subject is presented with three different stimuli. Each stimulus has a different feature along each one of the three feature dimensions (shape, color and texture). (B) The subject chooses one of the stimuli and receives a reward. (C) After a short delay a new trial begins.

Reinforcement Learning model

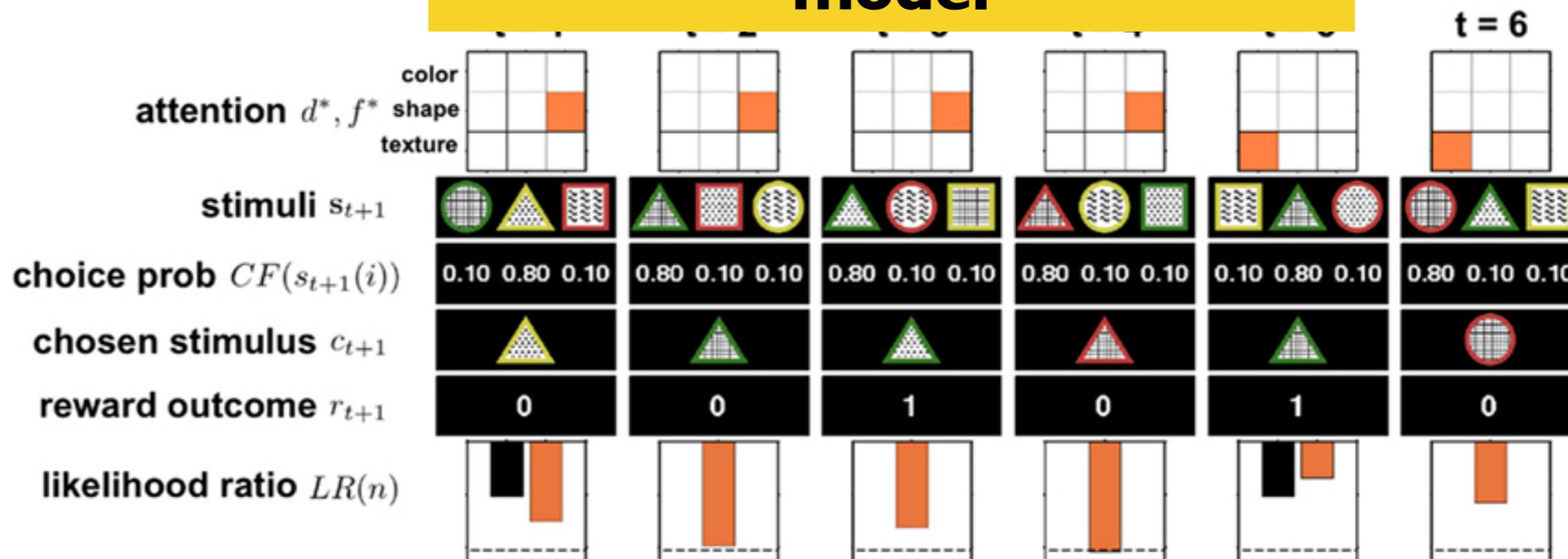


Figure 4: Illustration of six trials of the full selective attention model in action.

Example I: Self-Regulation Ontology

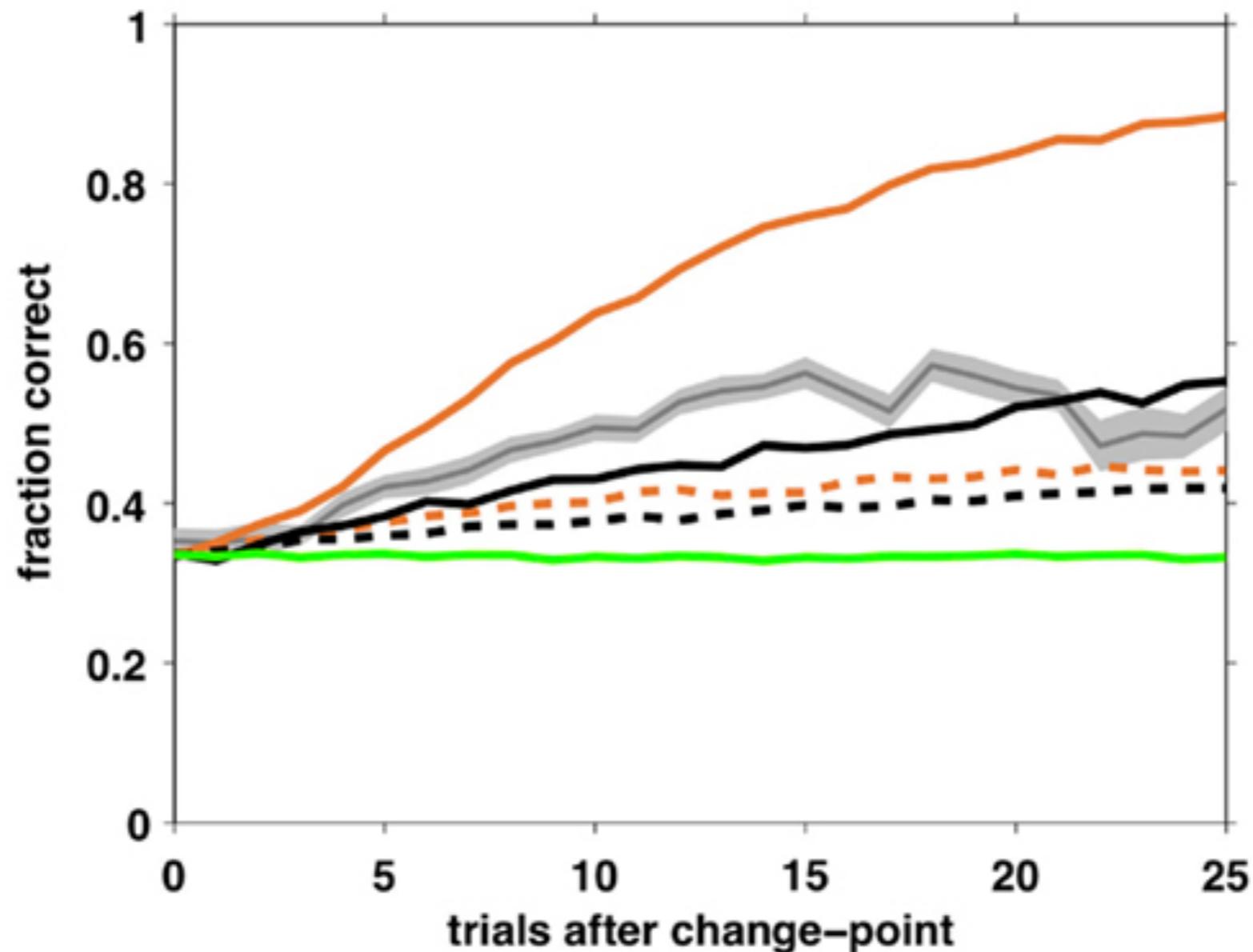
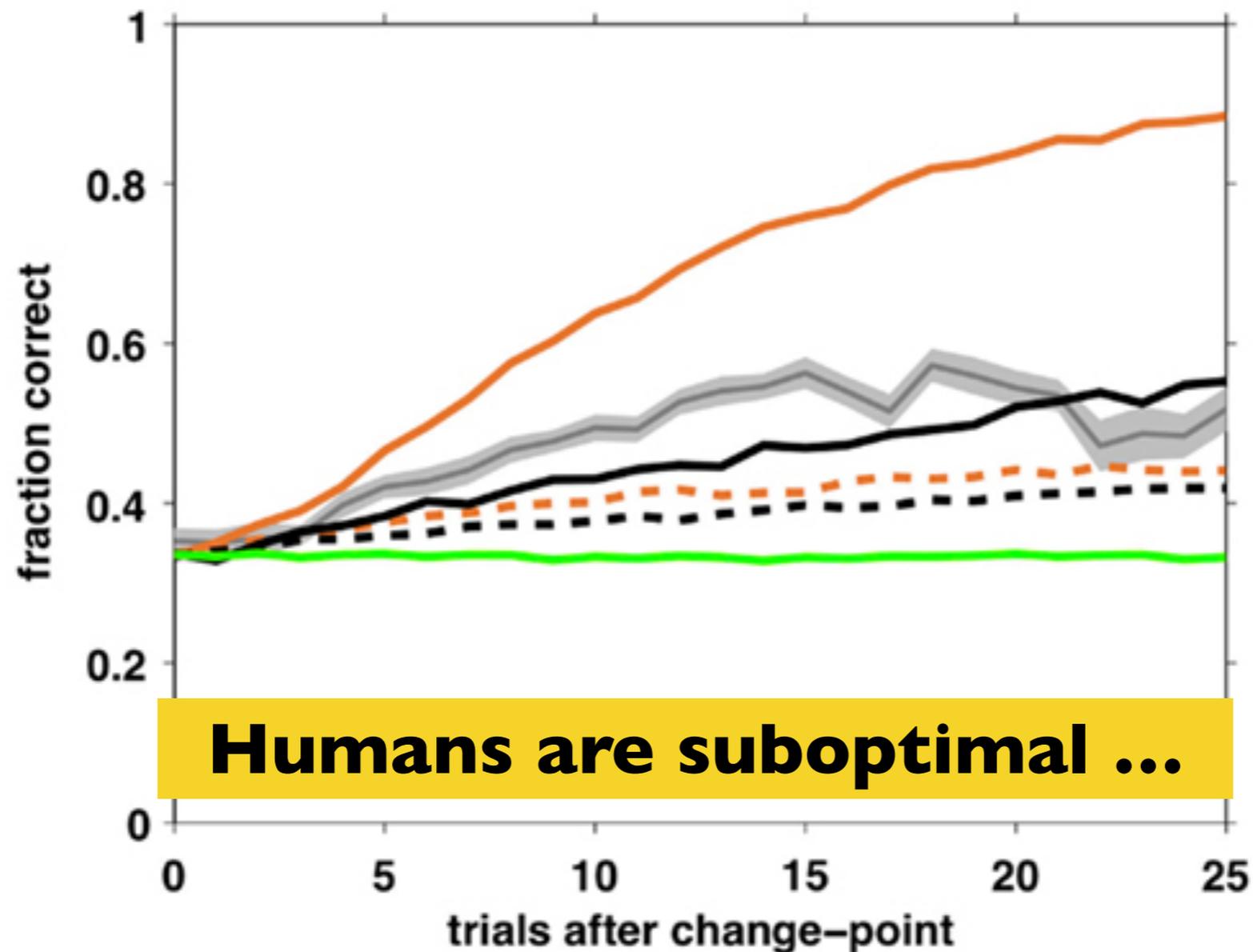


Figure 10. Average learning curves for simulated agents performing 10,000 trials of the task with un signaled changes. Black – full selective attention model with optimal (solid) and fit (dashed) parameters; Orange – Bayesian model with constant hazard rate with optimal (solid) and fit (dashed) parameters; Green – SA rand with fit parameter values; Gray – human data (shading: SEM).

Example I: Self-Regulation Ontology



Humans are suboptimal ...

Figure 10. Average learning curves for simulated agents performing 10,000 trials of the task with unsignaled changes. Black – full selective attention model with optimal (solid) and fit (dashed) parameters; Orange – Bayesian model with constant hazard rate with optimal (solid) and fit (dashed) parameters; Green – SA rand with fit parameter values; Gray – human data (shading: SEM).

Inferring relevance in a changing world

Robert C. Wilson* and Yael Niv

Department of Psychology, Neuroscience Institute, Princeton University, Princeton, NJ, USA

*“....From [these results], we conjecture that the problem of inferring relevance in general scenarios is too computationally demanding for the brain to solve optimally. **As a result the brain utilizes approximations...**”*

Table 2
Self-report surveys.

Self-Report Surveys	Individual Difference Variables	References
BIS-11	Attentional Motor	(Patton, Stanford, & Barratt, 1995)
BIS-BAS	Non-Planning BAS Drive BAS Fun-Seeking BAS Reward-Responsiveness BIS	(Carver & White, 1994)
Brief Self-Control Scale	Self-Control	(Roth, Isquith, & Gioia, 2005)
Dickman's Impulsivity Inventory	Dysfunctional Functional	(Dickman, 1990)
DOSPERT (EB/RP/RT)	Ethical Financial Health/Safety Recreational Social	(Blais & Weber, 2006)
Three-Factor Eating Questionnaire (R18)	Cognitive Restraint Emotional Eating Uncontrolled Eating	(de Lauzon et al., 2004)
Emotion Regulation Questionnaire	Reappraisal Suppression	(Gross & John, 2003)
Five Facet Mindfulness Questionnaire	Acts with Awareness Describe Non-Judgment Non-Reactive Observe	(Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006)
Future Time Perspective	Future-Time Perspective	(Carstensen & Lang, 1996)
Grit Scale	Grit	(Duckworth and Quinn, 2009)

Table 2
Self-report surveys.

Self-Report Surveys	Individual Difference Variables	References	Impulsive-Venturesome Survey	Impulsiveness Venturesomeness	
			Stanford Leisure-Time Activity <u>C</u> ategorical Item (L-Cat)	Activity Level	(Eysenck, Pearson, Easting, & Allsopp, 1985) (Kiernan et al., 2013)
BIS-11	Attentional Motor Non-Planning	(Patton, Stanford, & Barratt, 1995)	Mindful Attention Awareness Scale	Mindfulness	(Brown & Ryan, 2003)
BIS-BAS	BAS Drive BAS Fun-Seeking BAS Reward-Responsiveness	(Carver & White, 1994)	Multidimensional Personality Questionnaire (Control subscale)	Control	(Patrick, Curtin, & Tellegen, 2002)
Brief Self-Control Scale	BIS Self-Control	(Roth, Isquith, & Gioia, 2005)	Selection Optimization Compensation	Elective Selection Loss-based Selection Compensation Optimization	(Baltes, Baltes, Freund, & Lang, 1999)
Dickman's Impulsivity Inventory	Dysfunctional Functional	(Dickman, 1990)	Short Self-Regulation Survey	Control	(Carey, Neal, & Collins, 2004)
DOSPERT (EB/RP/RT)	Ethical Financial Health/Safety Recreational Social	(Blais & Weber, 2006)	Sensation Seeking Survey	Boredom Susceptibility Disinhibition Experience Seeking Thrill/Adventure Seeking	(Zuckerman, 2007)
Three-Factor Eating Questionnaire (R18)	Cognitive Restraint Emotional Eating Uncontrolled Eating	(de Lauzon et al., 2004)	Ten Item Personality Questionnaire	Agreeableness Conscientiousness Emotional Stability Extraversion Openness	(Gosling, Rentfrow, & Swann, 2003)
Emotion Regulation Questionnaire	Reappraisal Suppression	(Gross & John, 2003)	Theories of Willpower	Endorse Limited Resource	(Job, Dweck, & Walton, 2010)
Five Facet Mindfulness Questionnaire	Acts with Awareness Describe Non-Judgment Non-Reactive Observe	(Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006)	Time Perspective Survey	Future Past Negative Past Positive Present Fatalistic Present Hedonistic	(Zimbardo & Boyd, 2015)
Future Time Perspective	Future-Time Perspective	(Carstensen & Lang, 1996)	UPPS + P	Lack of Perseverance Lack of Premeditation Negative Urgency Positive Urgency Sensation Seeking	(Lynam, Smith, Whiteside, & Cyders, 2006)
Grit Scale	Grit	(Duckworth and Quinn, 2009)			

Impulsive-Venturesome Survey	Impulsiveness Venturesomeness	(Eysenck, Pearson, Easting, & Allsopp, 1985)
Stanford Leisure-Time Activity Categorical Item (L-Cat)	Activity Level	(Kiernan et al., 2013)
Mindful Attention Awareness Scale	Mindfulness	(Brown & Ryan, 2003)
Multidimensional Personality Questionnaire (Control subscale)	Control	(Patrick, Curtin, & Tellegen, 2002)
Selection Optimization Compensation	Elective Selection Loss-based Selection Compensation Optimization	(Baltes, Baltes, Freund, & Lang, 1999)
Short Self-Regulation Survey	Control	(Carey, Neal, & Collins, 2004)
Sensation Seeking Survey	Boredom Susceptibility Disinhibition Experience Seeking Thrill/Adventure Seeking	(Zuckerman, 2007)
Ten Item Personality Questionnaire	Agreeableness Conscientiousness Emotional Stability Extraversion Openness	(Gosling, Rentfrow, & Swann, 2003)
Theories of Willpower	Endorse Limited Resource	(Job, Dweck, & Walton, 2010)
Time Perspective Survey	Future Past Negative Past Positive Present Fatalistic Present Hedonistic	(Zimbardo & Boyd, 2015)
UPPS + P	Lack of Perseverance Lack of Premeditation Negative Urgency Positive Urgency Sensation Seeking	(Lynam, Smith, Whiteside, & Cyders, 2006)

Interest and Preference Test

Sensation Seeking Scale Form V

Directions: Each of the items below contains two choices, A and B. Please indicate (circle) on your answer sheet which of the choices most describes your likes or the way you feel. In some cases you may find items in which both choices describe your likes or feelings. Please choose the one which better describes your likes or feelings.

In some cases you may find items in which you do not like either choice. In these cases mark the choice you dislike least. Please try to answer each item.

It is important you respond to all items with only one choice, A or B. We are interested only in your likes or feeling, not in how others feel about these things or how one is supposed to feel. There are no right or wrong answers as in other kinds of tests. Be frank and give your honest appraisal of yourself.

1. A. I like "wild" uninhibited parties
B. I prefer quiet parties with good conversation
2. A. There are some movies I enjoy seeing a second or even a third time
B. I can't stand watching a movie that I've seen before
3. A. I often wish I could be a mountain climber
B. I can't understand people who risk their necks climbing mountains
4. A. I dislike all body odors
B. I like some for the earthly body smells
5. A. I get bored seeing the same old faces
B. I like to comfortable familiarity of everyday friends
6. A. I like to explore a strange city or section of town by myself, even if it means getting lost
B. I prefer a guide when I am in a place I don't know well
7. A. I dislike people who do or say things just to shock or upset others
B. When you can predict almost everything a person will do and say he or she must be a bore
8. A. I usually don't enjoy a movie or play where I can predict what will happen in advance
B. I don't mind watching a movie or a play where I can predict what will happen in advance
9. A. I have tried marijuana or would like to
B. I would never smoke marijuana

Impulsive-Venturesome Survey

Impulsiveness
Venturesomeness

Stanford Leisure-Time Activity Categorical Item (L-Cat)

Activity Level

Mindful Attention Awareness Scale

Mindfulness

Multidimensional Personality Questionnaire (Control subscale)

Control

Selection Optimization Compensation

Elective Selection
Loss-based Selection

Compensation
Optimization

Short Self-Regulation Survey

Control

Sensation Seeking Survey

Boredom Susceptibility
Disinhibition
Experience Seeking
Thrill/Adventure Seeking

Ten Item Personality Questionnaire

Agreeableness
Conscientiousness
Emotional Stability
Extraversion

Theories of Willpower

Openness
Endorse Limited Resource

Time Perspective Survey

Future
Past Negative
Past Positive
Present Fatalistic

UPPS + P

Present Hedonistic
Lack of Perseverance
Lack of Premeditation
Negative Urgency
Positive Urgency
Sensation Seeking

Example I: Self-Regulation Ontology

Demographics Variables

Sex	Age	Race	OtherRace	HispanicLatino	HighestEducation	HeightInches	WeightPounds	RelationshipStatus
DivorceCount	LongestRelationship	RelationshipNumber	ChildrenNumber	HouseholdIncome	RetirementAccount	RetirementPercentStocks	RentOwn	
MortgageDebt	CarDebt	EducationDebt	CreditCardDebt	OtherDebtSources			OtherDebtAmount	
CoffeeCupsPerDay	TeaCupsPerDay	CaffienatedSodaCansPerDay	CaffieneOtherSourcesDayMG	GamblingProblem	TrafficTicketsLastYearCount			
	TrafficAccidentsLifeCount	ArrestedChargedLifeCount	MotivationForParticipation					

Example I: Self-Regulation Ontology

Demographics Variables

Sex	Age	Race	OtherRace	HispanicLatino	HighestEducation	HeightInches	WeightPounds	RelationshipStatus
DivorceCount	LongestRelationship	RelationshipNumber	ChildrenNumber	HouseholdIncome	RetirementAccount	RetirementPercentStocks	RentOwn	
MortgageDebt	CarDebt	EducationDebt	CreditCardDebt	OtherDebtSources			OtherDebtAmount	
CoffeeCupsPerDay	TeaCupsPerDay	CaffienatedSodaCansPerDay	CaffieneOtherSourcesDayMG	GamblingProblem	TrafficTicketsLastYearCount			
TrafficAccidentsLifeCount	ArrestedChargedLifeCount	MotivationForParticipation						

... plus alcohol/drugs variables

LifetimeSmoke100Cigs	HowLongSmoked	SmokeEveryDay	CigsPerDay	HowSoonSmokeAfterWaking	OtherTobaccoProducts	AlcoholHowOften		
HowOftenGuiltRemorseDrinking	HowOftenUnableRememberDrinking	InjuredDrinking	RelativeFriendConcernedDrinking	CannabisPast6Months				
AlcoholHowManyDrinksDay	AlcoholHowOften6Drinks	HowOftenCantStopDrinking	HowOftenFailedActivitiesDrinking	HowOftenDrinkMorning				
CannabisHowOften	CannabisHoursStoned	HowOftenCantStopCannabis	HowOftenFailedActivitiesCannabis	HowOftenDevotedTimeCannabis				
HowOftenMemoryConcentrationProblemCannabis	HowOftenHazardousCannabis	CannabisConsideredReduction	OtherDrugs	AbuseMoreThanOneDrugAtA				
AbleToStopDrugs	BlackoutFlashbackDrugUse	FeelBadGuiltyDrugUse	SpouseParentsComplainDrugUse	NeglectedFamilyDrugUse				
EngagedInIllegalActsToObtainDrugs	WithdrawalSymptoms	MedicalProblemsDueToDrugUse						

Example I: Self-Regulation Ontology

Health Targets

Nervous	Hopeless	RestlessFidgety	Depressed	EverythingIsEffort	Worthless	Last30DaysUsual	DaysLostLastMonth	DaysHalfLastMonth
---------	----------	-----------------	-----------	--------------------	-----------	-----------------	-------------------	-------------------

DoctorVisitsLastMonth	DaysPhysicalHealthFeelings	PsychDiagnoses	PsychDiagnosesOther
-----------------------	----------------------------	----------------	---------------------

NeurologicalDiagnoses	NeurologicalDiagnosesDescribe
-----------------------	-------------------------------

Health outcomes

Nervous Hopeless RestlessFidgety Depressed EverythingIsEffort Worthless Last30DaysUsual DaysLostLastMonth DaysHalfLastMonth

DoctorVisitsLastMonth DaysPhysicalHealthFeelings PsychDiagnoses PsychDiagnosesDescribe

NeurologicalDiagnoses NeurologicalDiagnosesDescribe

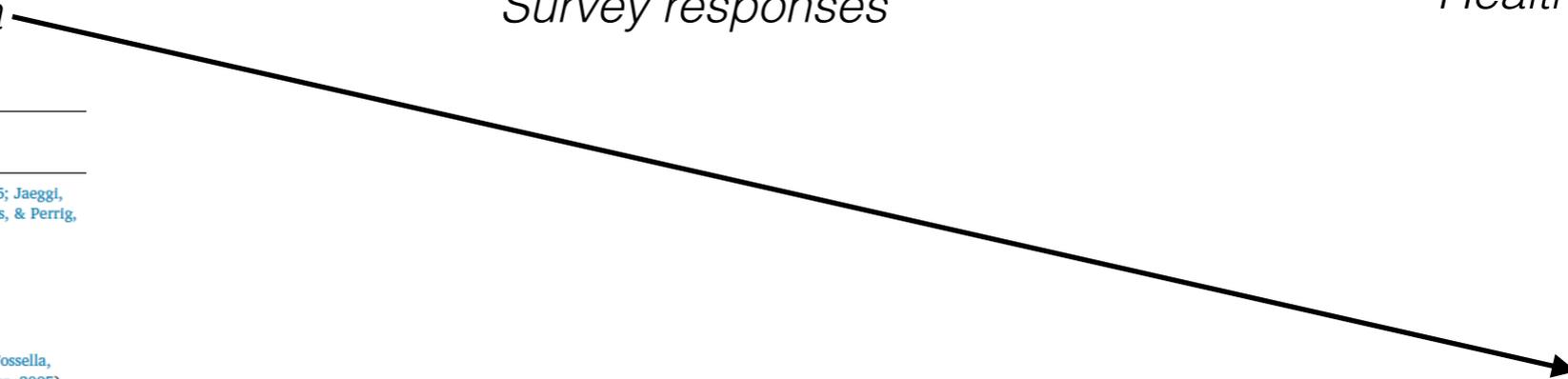
Task performance data

Survey responses

Health outcomes

Table 1
Cognitive tasks.

Task	Individual Difference Variables	References
Adaptive N-Back	DDM Parameters Average load	(Harvey et al., 2005; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008)
Angling Risk Task	Two Conditions (Keep, Release): Adjusted Clicks Loss Percent Score	(Pleskac, 2008)
Attention Network Task	DDM Parameters Alerting Effect Orient Effect Conflict Effect	(Fan, McCandliss, Fossella, Flombaum, & Posner, 2005)
Bickel Titrator	Discount Rate for three payout magnitudes	(Koffarnus & Bickel, 2014)
Choice Reaction Time Cognitive Reflection Task	DDM Parameters Correct Proportion Intuitive Proportion	(Primi, Morsanyi, Chiesi, Donati, & Hamilton, 2016; Toplak, West, & Stanovich, 2014)
Columbia Card Task Cold/Hot	Average # of cards chosen Gain Sensitivity Loss Sensitivity # Loss Cards Sensitivity Level of Information Use	(Figner, Mackinlay, Wilkening, & Weber, 2009)
Dietary Decision Task	Health Sensitivity Taste Sensitivity	(Hare, Camerer, & Rangel, 2009)
Digit Span	Forward Span Reverse Span	(Woods et al., 2011)
Directed Forgetting	DDM Parameters Proactive Interference	(Nee, Jonides, & Berman, 2007)
Discount Titrator	Percent Patient	(Figner et al., 2010)
Dot Pattern Expectancy	DDM Parameters AY-BY BX-BY D-prime Bias	(Otto, Skatova, Madlon-Kay, & Daw, 2013)



NeurologicalDiagnoses	DoctorVisitsLastMonth	Nervous
	DaysPhysicalHealthFeelings	Hopeless
PsychDiagnoses	DaysPhysicalHealthFeelings	RestlessFidgety
	PsychDiagnoses	Depressed
NeurologicalDiagnosesDescribe	PsychDiagnoses	EverythingIsEffort
	PsychDiagnosesDescribe	Worthless
PsychDiagnosesOther	PsychDiagnosesDescribe	Last30DaysUsual
	PsychDiagnosesOther	DaysLostLastMonth
		DaysHalfLastMonth

Task performance data

Survey responses

Health outcomes

Table 1
Cognitive tasks.

Task	Individual Difference Variables	References
Adaptive N-Back	DDM Parameters Average load	(Harvey et al., 2005; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008)
Angling Risk Task	Two Conditions (Keep, Release): Adjusted Clicks Loss Percent Score	(Pleskac, 2008)
Attention Network Task	DDM Parameters Alerting Effect Orient Effect Conflict Effect	(Fan, McCandliss, Fossella, Flombaum, & Posner, 2005)
Bickel Titrator	Discount Rate for three payout magnitudes	(Koffarnus & Bickel, 2014)
Choice Reaction Time Cognitive Reflection Task	DDM Parameters Correct Proportion Intuitive Proportion	(Primi, Morsanyi, Chiesi, Donati, & Hamilton, 2016; Toplak, West, & Stanovich, 2014)
Columbia Card Task Cold/Hot	Average # of cards chosen Gain Sensitivity Loss Sensitivity # Loss Cards Sensitivity Level of Information Use	(Figner, Mackinlay, Wilkening, & Weber, 2009)
Dietary Decision Task	Health Sensitivity Taste Sensitivity	(Hare, Camerer, & Rangel, 2009)
Digit Span	Forward Span Reverse Span	(Woods et al., 2011)
Directed Forgetting	DDM Parameters Proactive Interference	(Nee, Jonides, & Berman, 2007)
Discount Titrator	Percent Patient	(Figner et al., 2010)
Dot Pattern Expectancy	DDM Parameters AY-BY BX-BY D-prime Bias	(Otto, Skatova, Madlon-Kay, & Daw, 2013)

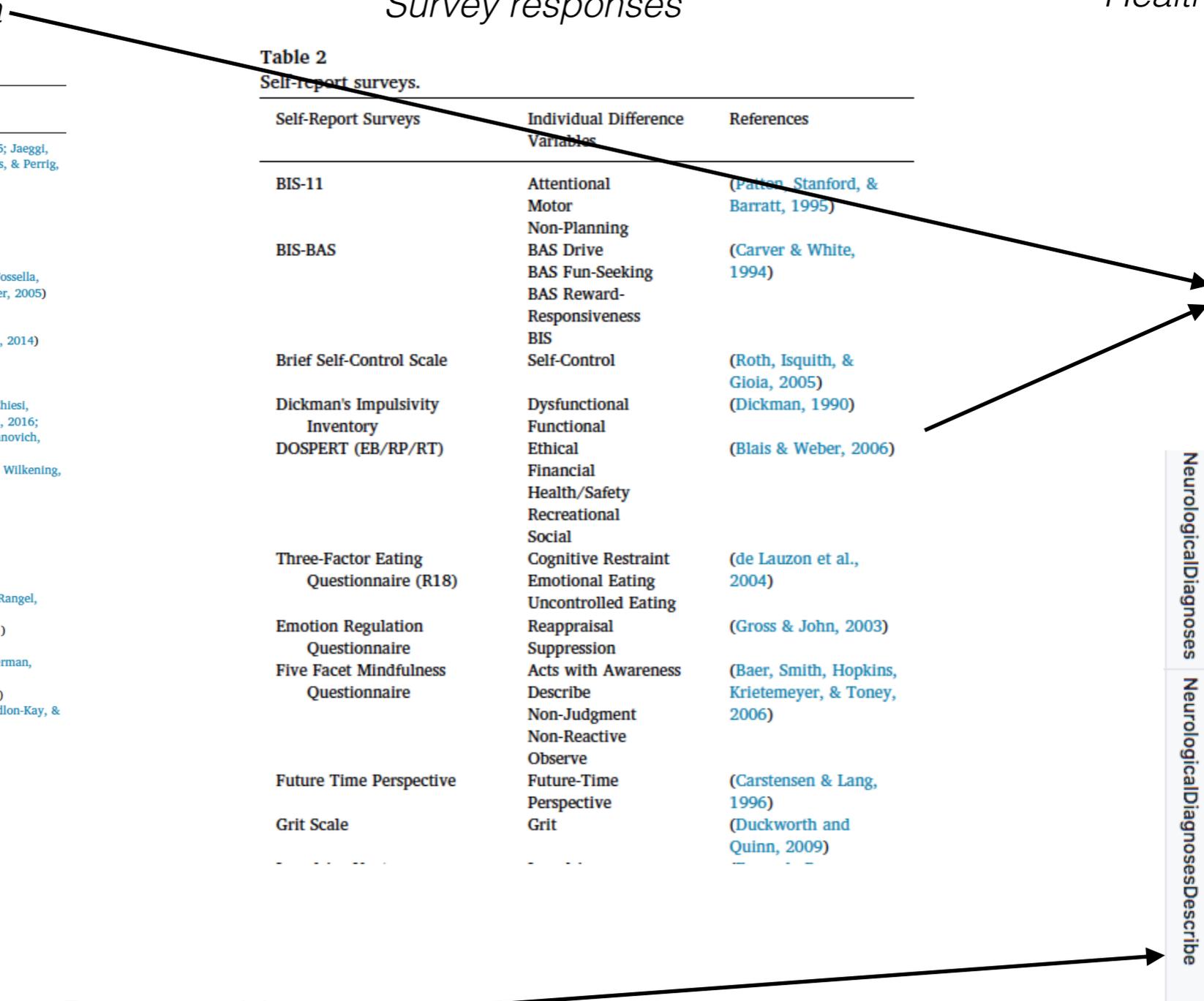
Table 2
Self-report surveys.

Self-Report Surveys	Individual Difference Variables	References
BIS-11	Attentional Motor Non-Planning	(Patten, Stanford, & Barratt, 1995)
BIS-BAS	BAS Drive BAS Fun-Seeking BAS Reward-Responsiveness BIS	(Carver & White, 1994)
Brief Self-Control Scale	Self-Control	(Roth, Isquith, & Gioia, 2005)
Dickman's Impulsivity Inventory DOSPERT (EB/RP/RT)	Dysfunctional Functional Ethical Financial Health/Safety Recreational Social	(Dickman, 1990) (Blais & Weber, 2006)
Three-Factor Eating Questionnaire (R18)	Cognitive Restraint Emotional Eating Uncontrolled Eating	(de Lauzon et al., 2004)
Emotion Regulation Questionnaire	Reappraisal Suppression	(Gross & John, 2003)
Five Facet Mindfulness Questionnaire	Acts with Awareness Describe Non-Judgment Non-Reactive Observe	(Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006)
Future Time Perspective	Future-Time Perspective	(Carstensen & Lang, 1996)
Grit Scale	Grit	(Duckworth and Quinn, 2009)

Demographics

Sex	Age	Race	OtherRace	HispanicLatino	HighestEducation	HeightInches	WeightPounds	RelationshipStatus
DivorceCount	LongestRelationship	RelationshipNumber	ChildrenNumber	HouseholdIncome	RetirementAccount	RetirementPercentStocks	RentOwn	
MortgageDebt	CarDebt	EducationDebt	CreditCardDebt	OtherDebtSources		OtherDebtAmount		
CoffeeCupsPerDay	TeaCupsPerDay	CaffienatedSodaCansPerDay	CaffieneOtherSourcesDayMG	GamblingProblem	TrafficTicketsLastYearCount			
TrafficAccidentsLifeCount	ArrestedChargedLifeCount	MotivationForParticipation						

DoctorVisitsLastMonth	Nervous
DaysPhysicalHealthFeelings	Hopeless
PsychDiagnoses	RestlessFidgety
NeurologicalDiagnoses	Depressed
PsychDiagnoses	EverythingIsEffort
NeurologicalDiagnosesDescribe	Worthless
PsychDiagnosesOther	Last30DaysUsual
	DaysLostLastMonth
	DaysHalfLastMonth



Task performance data

Survey responses

Health outcomes

Table 1
Cognitive tasks.

Task	Individual Difference Variables	References
Adaptive N-Back	DDM Parameters Average load	(Harvey et al., 2005; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008)
Angling Risk Task	Two Conditions (Keep, Release): Adjusted Clicks Loss Percent Score	(Pleskac, 2008)
Attention Network Task	DDM Parameters Alerting Effect Orient Effect Conflict Effect	(Fan, McCandliss, Fossella, Flombaum, & Posner, 2005)
Bickel Titrator	Discount Rate for three payout magnitudes	(Koffarnus & Bickel, 2014)
Choice Reaction Time Cognitive Reflection Task	DDM Parameters Correct Proportion Intuitive Proportion	(Primi, Morsanyi, Chiesi, Donati, & Hamilton, 2016; Toplak, West, & Stanovich, 2014)
Columbia Card Task Cold/Hot	Average # of cards chosen Gain Sensitivity Loss Sensitivity # Loss Cards Sensitivity Level of Information Use	(Figner, Mackinlay, Wilkening, & Weber, 2009)
Dietary Decision Task	Health Sensitivity Taste Sensitivity	(Hare, Camerer, & Rangel, 2009)
Digit Span	Forward Span Reverse Span	(Woods et al., 2011)
Directed Forgetting	DDM Parameters Proactive Interference	(Nee, Jonides, & Berman, 2007)
Discount Titrator	Percent Patient	(Figner et al., 2010)
Dot Pattern Expectancy	DDM Parameters AY-BY BX-BY D-prime Bias	(Otto, Skatova, Madlon-Kay, & Daw, 2013)

Table 2
Self-report surveys.

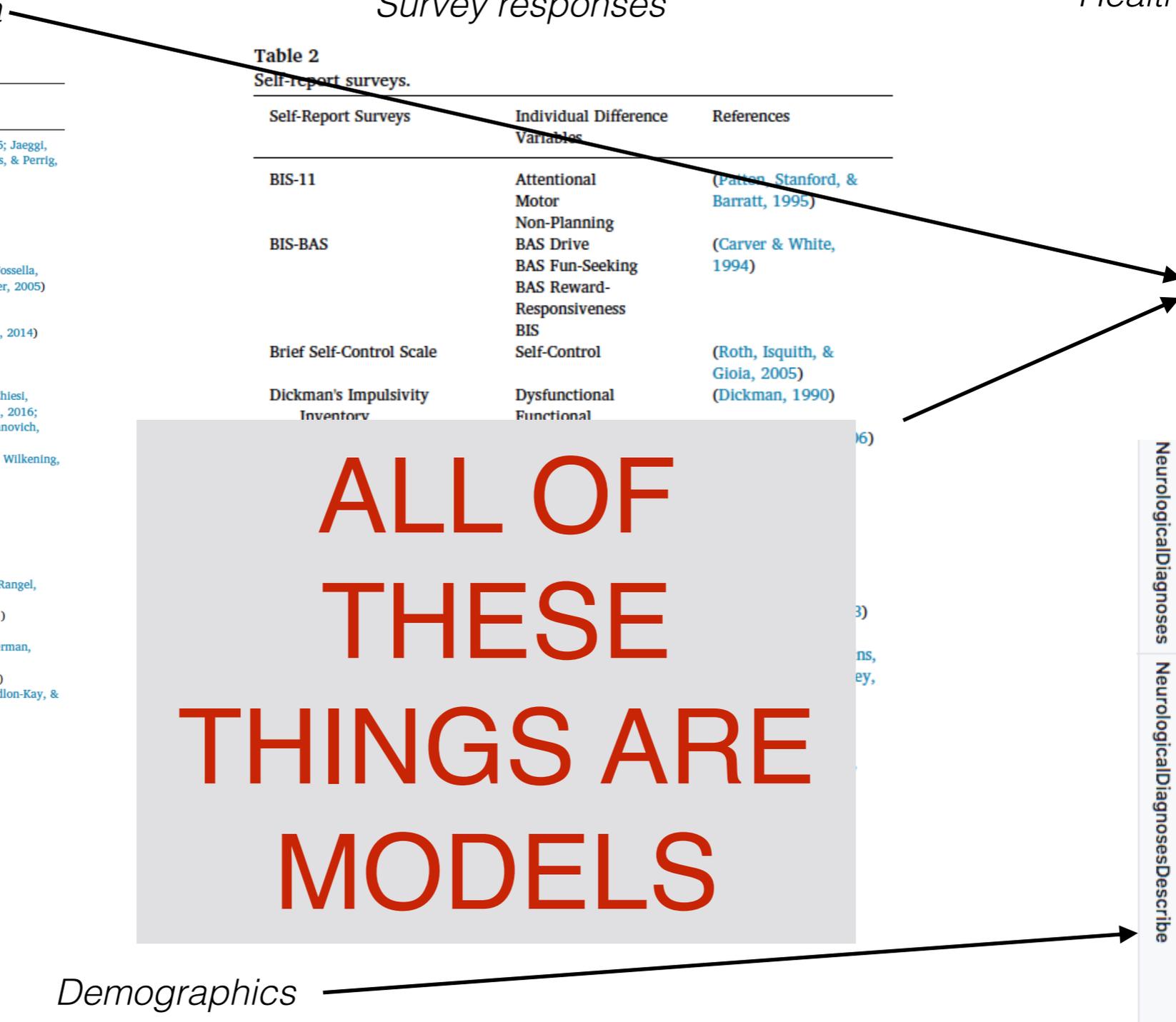
Self-Report Surveys	Individual Difference Variables	References
BIS-11	Attentional Motor Non-Planning	(Patten, Stanford, & Barratt, 1995)
BIS-BAS	BAS Drive BAS Fun-Seeking BAS Reward-Responsiveness BIS	(Carver & White, 1994)
Brief Self-Control Scale	Self-Control	(Roth, Isquith, & Gioia, 2005)
Dickman's Impulsivity Inventory	Dysfunctional Functional	(Dickman, 1990)

ALL OF THESE THINGS ARE MODELS

Demographics

Sex	Age	Race	OtherRace	HispanicLatino	HighestEducation	HeightInches	WeightPounds	RelationshipStatus
DivorceCount	LongestRelationship	RelationshipNumber	ChildrenNumber	HouseholdIncome	RetirementAccount	RetirementPercentStocks	RentOwn	
MortgageDebt	CarDebt	EducationDebt	CreditCardDebt	OtherDebtSources		OtherDebtAmount		
CoffeeCupsPerDay	TeaCupsPerDay	CaffienatedSodaCansPerDay	CaffieneOtherSourcesDayMG	GamblingProblem	TrafficTicketsLastYearCount			
TrafficAccidentsLifeCount	ArrestedChargedLifeCount	MotivationForParticipation						

DoctorVisitsLastMonth	Nervous
DaysPhysicalHealthFeelings	Hopeless
PsychDiagnoses	RestlessFidgety
NeurologicalDiagnoses	Depressed
NeurologicalDiagnosesDescribe	EverythingIsEffort
PsychDiagnosesOther	Worthless
	Last30DaysUsual
	DaysLostLastMonth
	DaysHalfLastMonth



Example I: Self-Regulation Ontology

https://github.com/IanEisenberg/Self_Regulation_Ontology

ianEisenberg / Self_Regulation_Ontology

Watch 1 Star 0 Fork 0

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights

Branch: master

Self_Regulation_Ontology / Data / Complete_02-16-2019 /

Create new file Upload files Find file History

ianEisenberg first commit for new SRO repository. Removes history kept in Self_Reg... Latest commit b8f721e 28 days ago

..		
Individual_Measures	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
metadata	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
references	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
DV_valence.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
README.txt	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
alcohol_drugs.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
alcohol_drugs_ordinal.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
demographic_health.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
demographics.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
demographics_ordinal.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
health.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
health_ordinal.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
items.csv.gz	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
meaningful_variables.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
meaningful_variables_EZ.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago

Example I: Self-Regulation Ontology

https://github.com/IanEisenberg/Self_Regulation_Ontology

The screenshot displays the GitHub repository page for 'Self_Regulation_Ontology' by IanEisenberg. The repository is on the 'master' branch. The file list shows the following structure:

File/Folder	Commit Message	Time
..		
Individual_Measures	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
metadata	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
references	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
DV_valence.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
README.txt	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
alcohol_drugs.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
alcohol_drugs_ordinal.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
demographic_health.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
demographics.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
demographics_ordinal.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
health.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
health_ordinal.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
items.csv.gz	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
meaningful_variables.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago
meaningful_variables_EZ.csv	first commit for new SRO repository. Removes history kept in Self_Reg...	28 days ago

https://github.com/IanEisenberg/Self_Regulation_Ontology/blob/master/Data/Complete_02-16-2019/README.txt

```
1 demographics_survey.csv: demographic information from expfactory-surveys
2
3 alcohol_drug_survey.csv: alcohol, smoking, marijuana and other drugs from expfactory-surveys
4
5 ky_survey.csv: mental health and neurological/health conditions from expfactory-surveys
6
7 items.csv.gz: gzipped csv of all item information across surveys
8
9 subject_x_items.csv: reshaped items.csv such that rows are subjects and columns are individual items
10
11 Individual Measures: directory containing gzip compressed files for each individual measures
12
13 DV_valence.csv: Subjective assessment of whether each variable's 'natural' direction implies 'better' self regulation
14
15 variables_exhaustive.csv: all variables calculated for each measure
16
17 meaningful_variables_noDDM.csv: subset of exhaustive data to only meaningful variables with DDM parameters removed
18
19 meaningful_variables_EZ.csv: subset of exhaustive data to only meaningful variables with rt/acc parameters removed (replaced by
20
21 meaningful_variables_hddm.csv: subset of exhaustive data to only meaningful variables with rt/acc parameters removed (replaced
22
23 meaningful_variables.csv: Same as meaningful_variables_hddm.csv
24
25 meaningful_variables_clean.csv: same as meaningful_variables.csv with skewed variables transformed and then outliers removed
26
27 meaningful_variables_imputed.csv: meaningful_variables_clean.csv after imputation with missForest
28
29 taskdata*.csv: taskdata are the same as meaningful_variables excluded surveys. Note that imputation is performed on the entire
30
31 short*.csv: short versions are the same as long versions with variable names shortened using variable_name_lookup.csv
32
```

Example I: Self-Regulation Ontology

[IPYNB: Self-Regulation Ontology data]

Example 2: The Ventral Visual Pathway

Understanding complex, noisy data streams is a critical part of cognition.

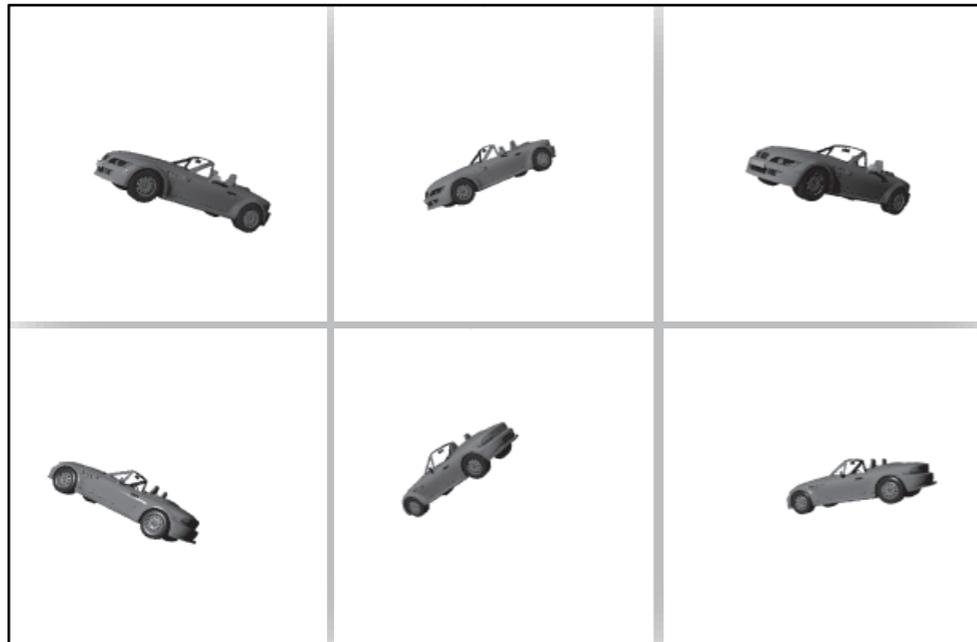


“Mercedes behind Lamborghini, on a field in front of mountains.”

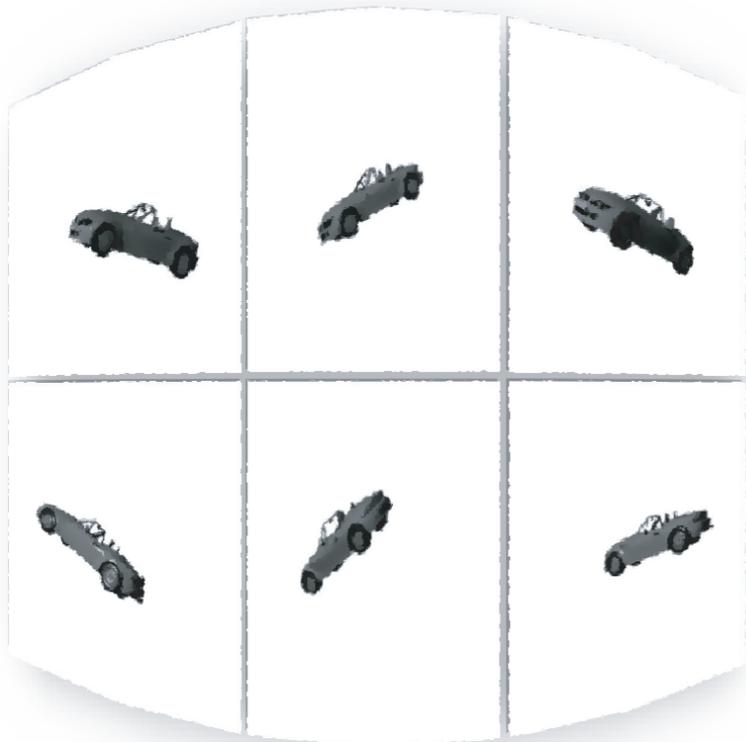
Example 2: The Ventral Visual Pathway



View: position, size, pose, illumination



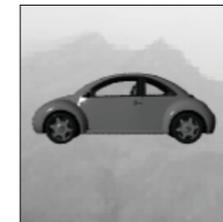
Distortion & Noise



Background variation

Geometric variation

Beetle



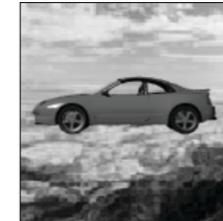
BMW Z3



Clio



Celica



Alfa



car identities

VW Bora



BMW 325

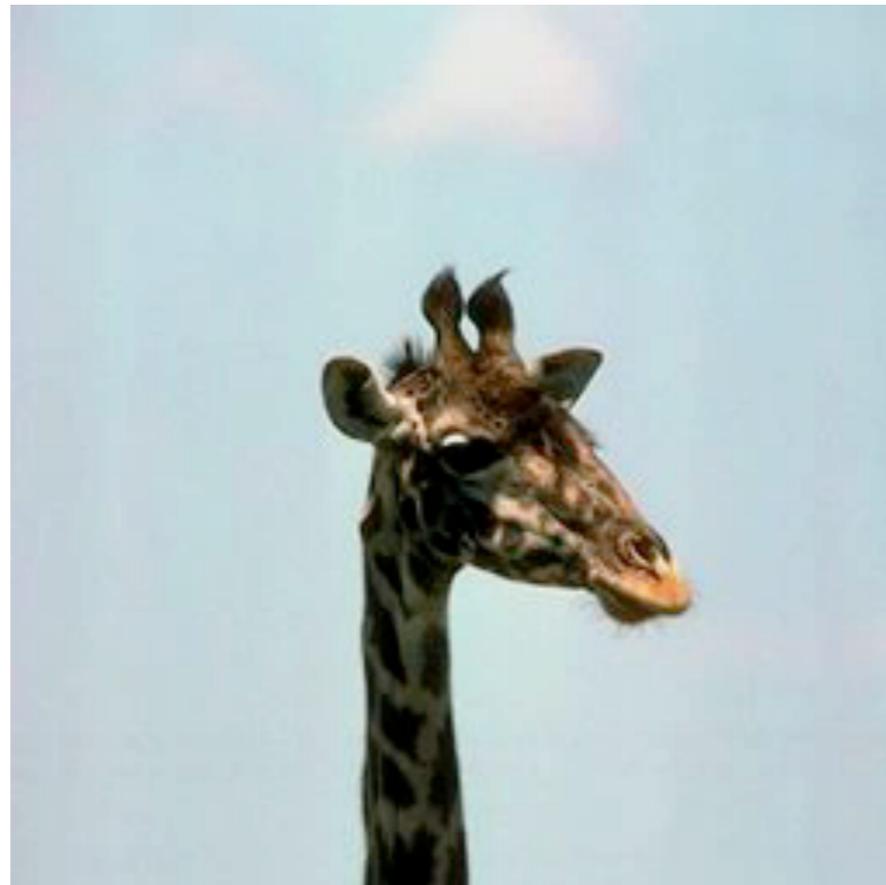


Astra



Example 2: The Ventral Visual Pathway

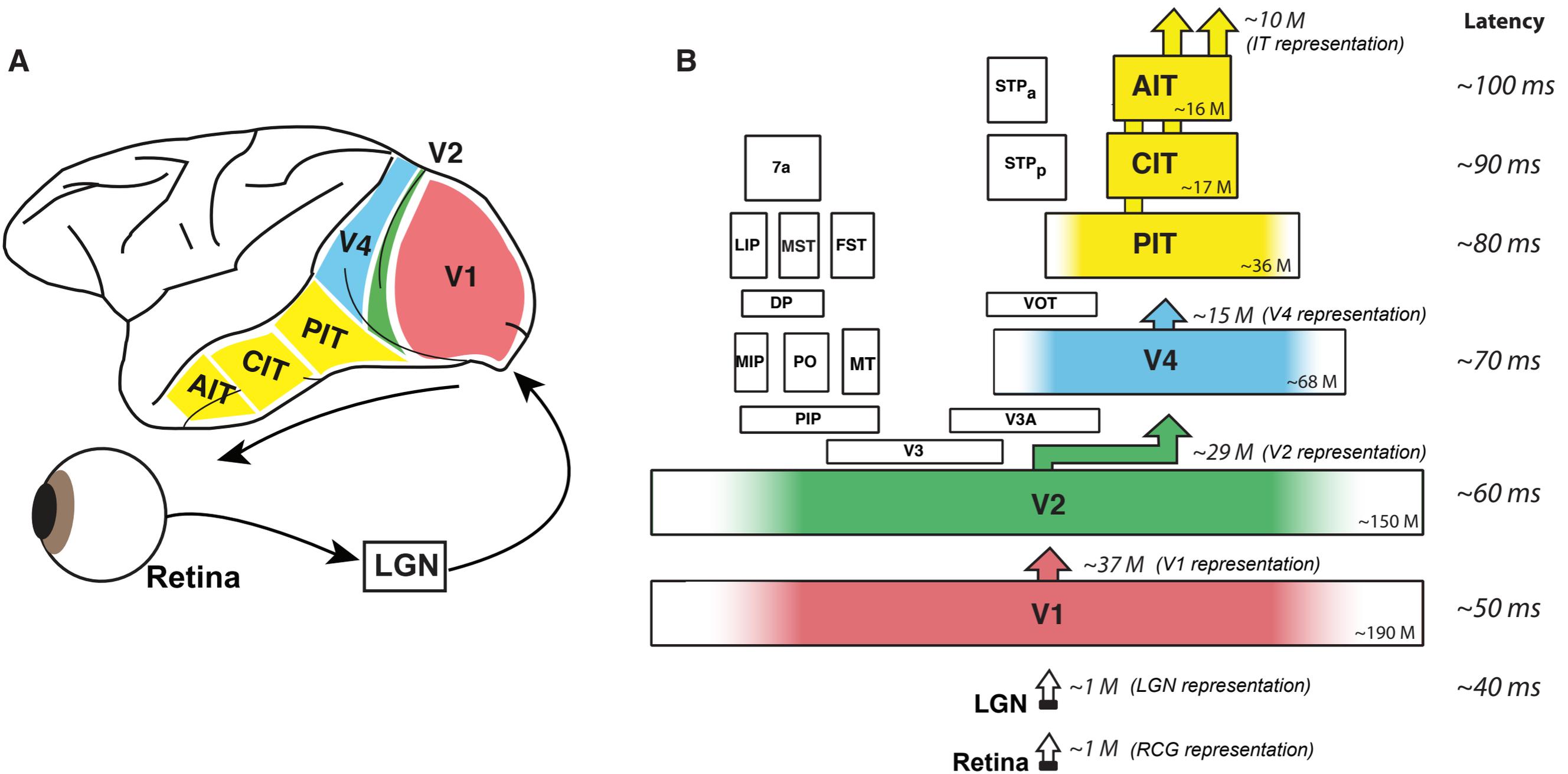
Understanding complex, noisy data streams is a critical part of cognition.



And you can do it fast.

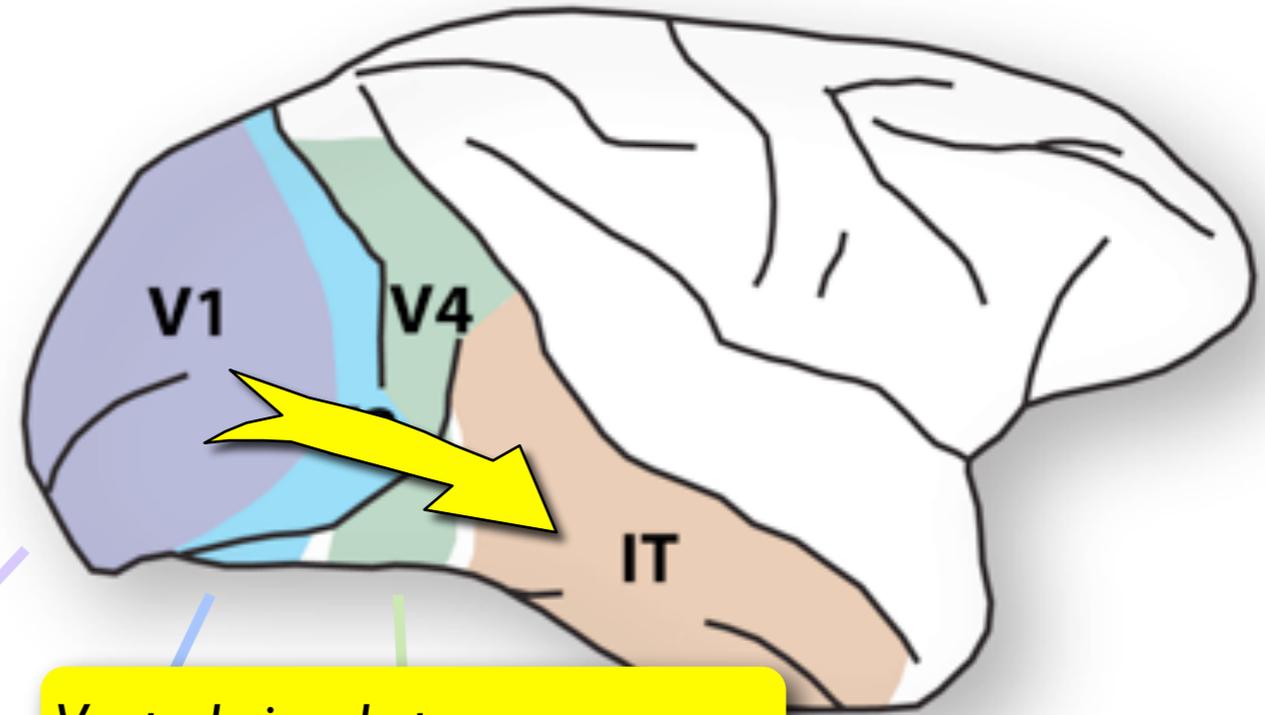
Example 2: The Ventral Visual Pathway

Neuroanatomical, cytoarchitectonic, and latency evidence:

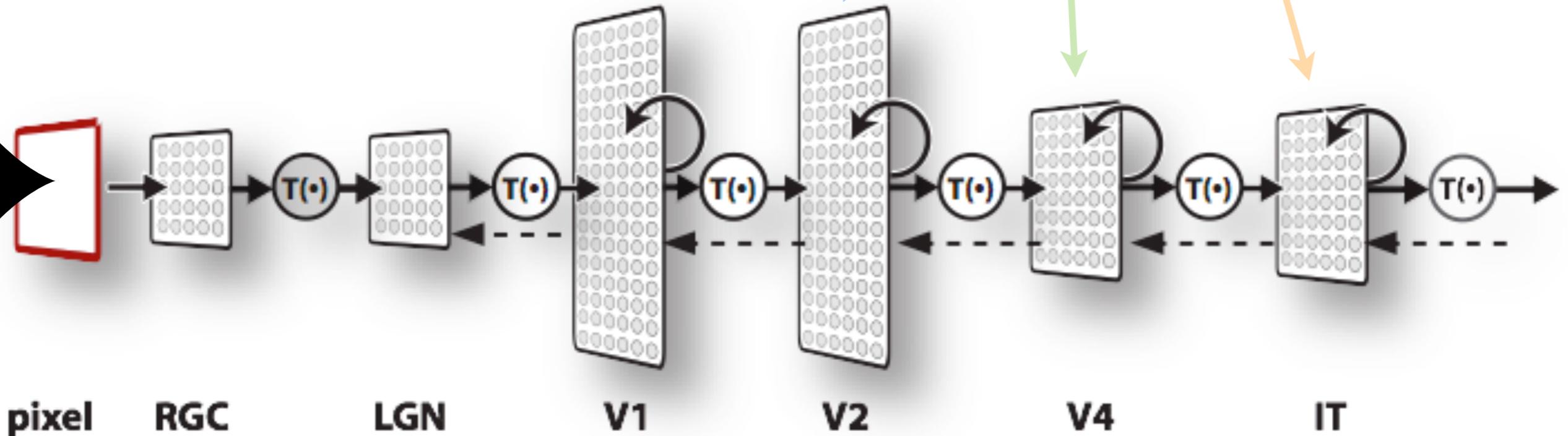


Example 2: The Ventral Visual Pathway

Kaas (2003), Van Essen (2003), Valois and Morgan (1974), Gross (1973), Mishkin and Ungerleider (1983), Holmes and Gross (1984), Harel et al. (1987), Freiwald and Tsao (2010), Pitcher, et al. (2009), Yaginuma (1982), Holmes (1984), Weiskrantz (1984), Schiller (1995), Afraz (2006), Verhoef (2012), Rust (2010), Freiwald (2010), Lehky (2007), Majaj (2015)



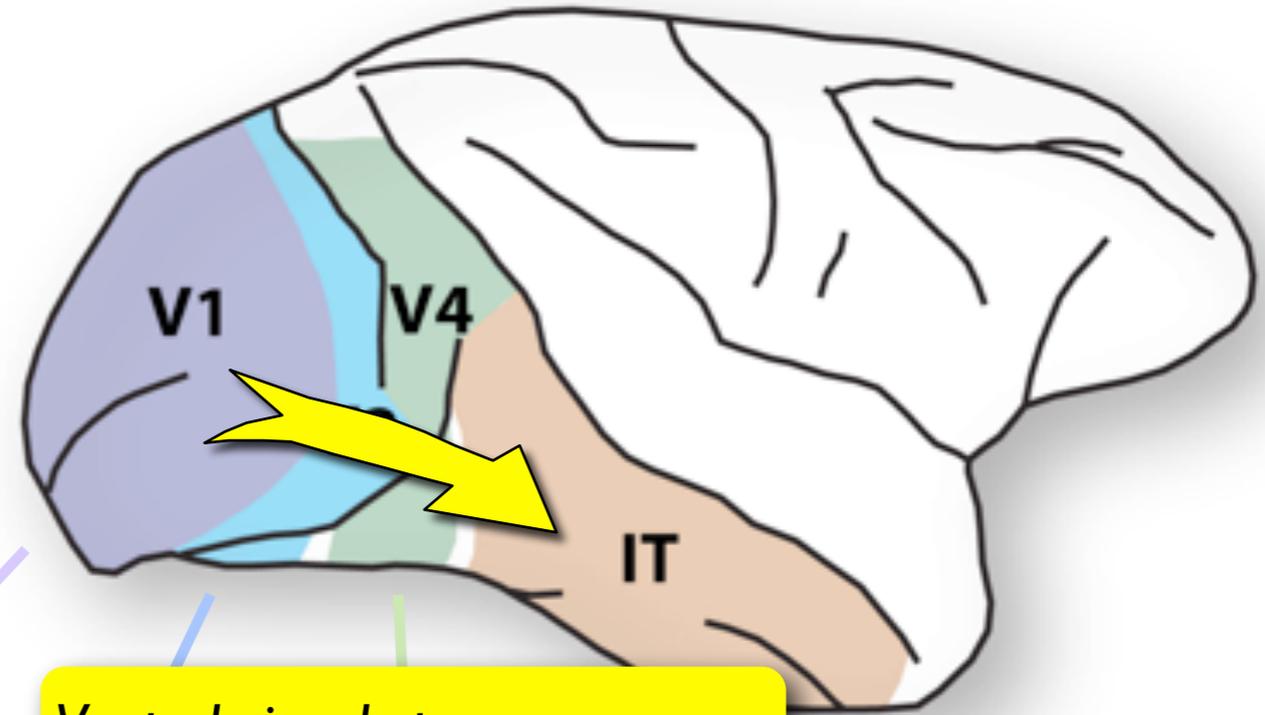
Ventral visual stream



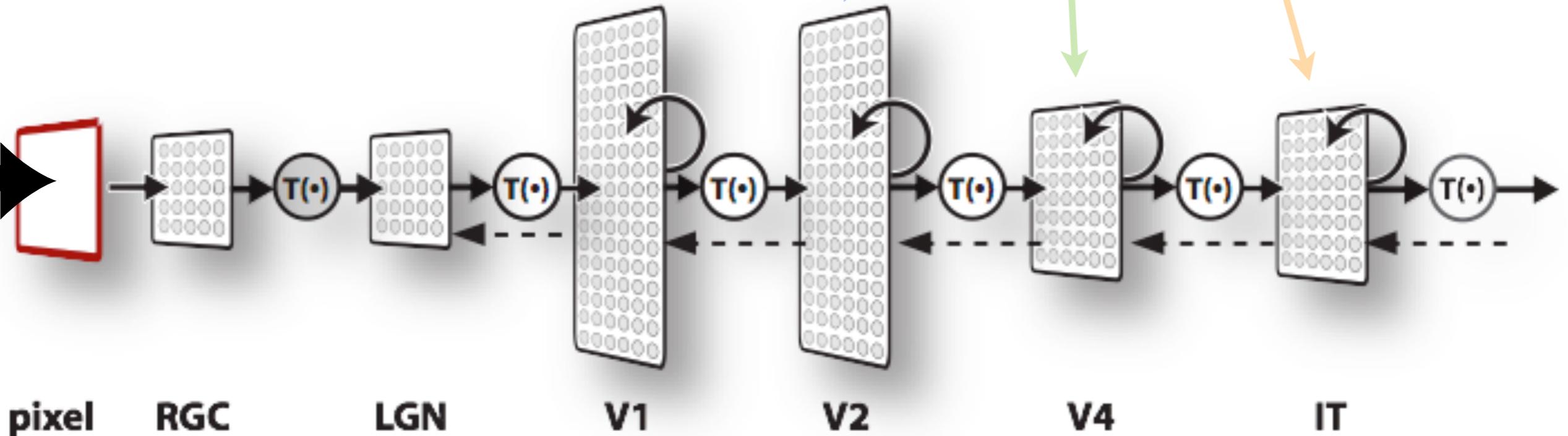
Example 2: The Ventral Visual Pathway



rhesus macaque (macaca mulatta)



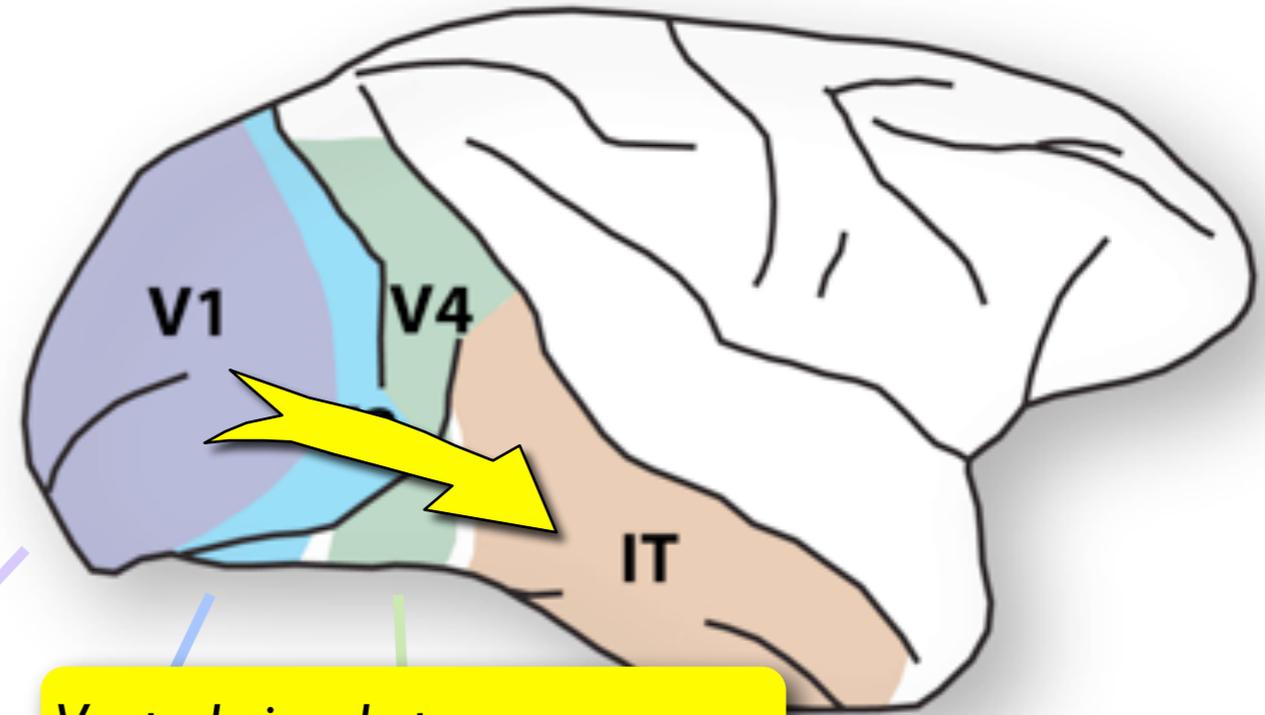
Ventral visual stream



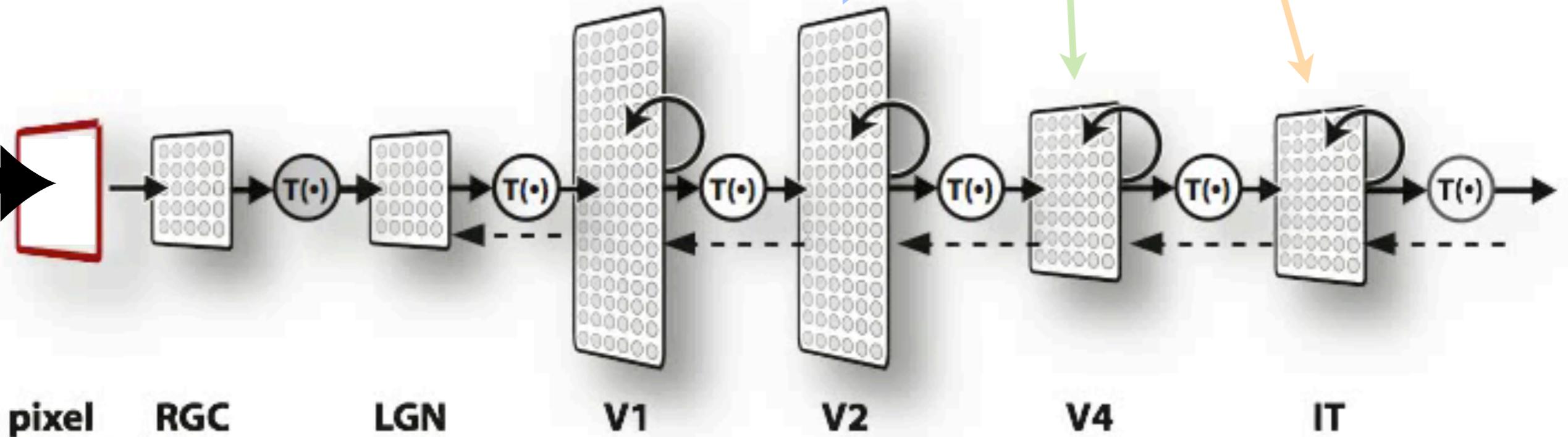
Example 2: The Ventral Visual Pathway



rhesus macaque (macaca mulatta)

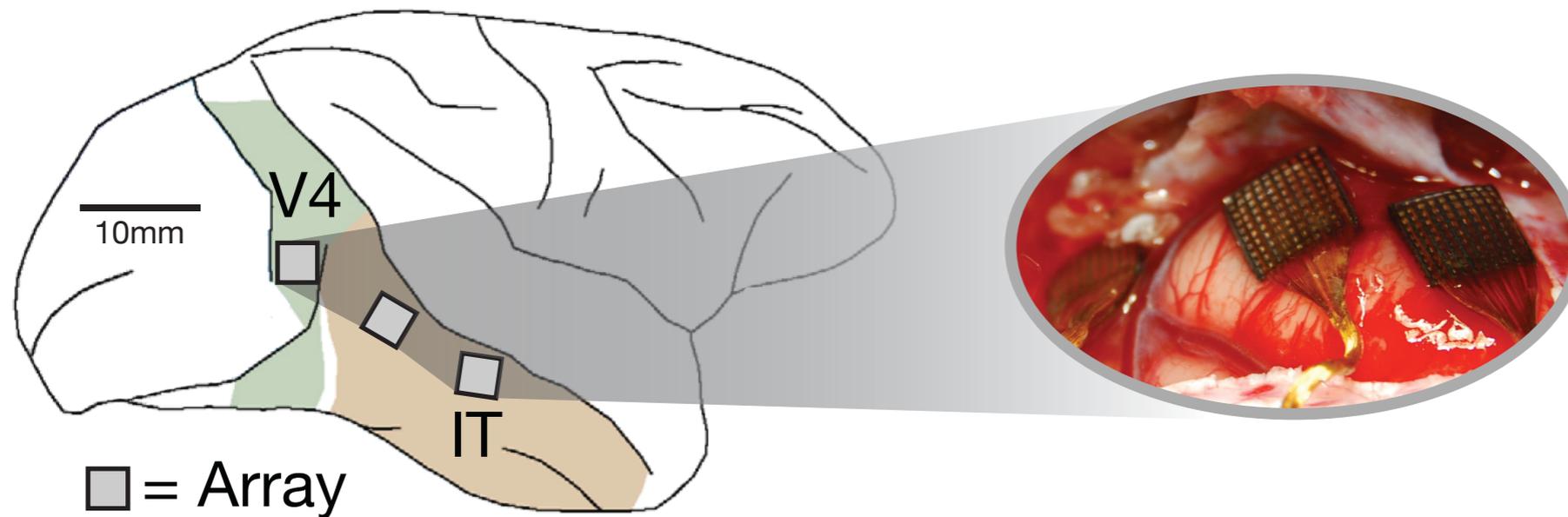


Ventral visual stream

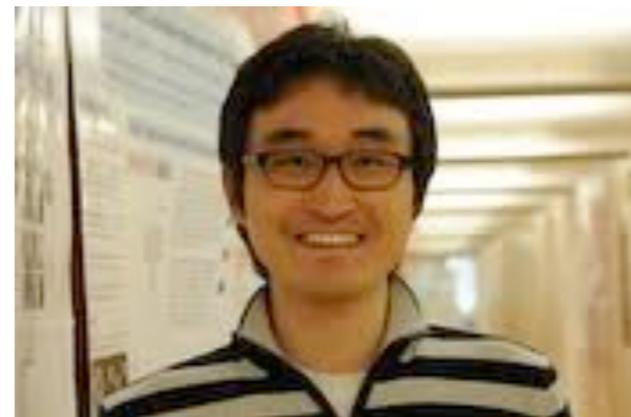


Example 2: The Ventral Visual Pathway

Multi-array electrophysiology in macaque V4 and IT.



About 300 total sites



Ha Hong



Jim DiCarlo

Example 2: The Ventral Visual Pathway

Ventral Cortex Data: from multi-array electrophysiology in macaques

Majaj, Najib J., Ha Hong, Ethan A. Solomon, and James J. DiCarlo. "Simple learned weighted sums of inferior temporal neuronal firing rates accurately predict human core object recognition performance." *Journal of Neuroscience* 35, no. 39 (2015): 13402-13418.

First paper comparing deep neural networks to visual cortex: Yamins Daniel LK, Ha Hong, Charles F. Cadieu, Ethan A. Solomon, Darren Seibert, and James J. DiCarlo. "Performance-optimized hierarchical models predict neural responses in higher visual cortex." *Proceedings of the National Academy of Sciences* 111, no. 23 (2014): 8619-8624.

Hong, Ha, Daniel LK Yamins, Najib J. Majaj, and James J. DiCarlo. "Explicit information for category-orthogonal object properties increases along the ventral stream." *Nature neuroscience* 19, no. 4 (2016): 613.

Example 2: The Ventral Visual Pathway

5760 images

64 objects

8 categories

uncorrelated photo backgrounds

Low variation



... 640 images

Medium variation



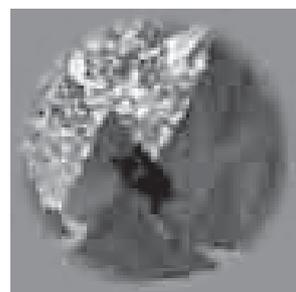
... 2560 images

High variation



... 2560 images

Animals



Boats



Cars



Chairs



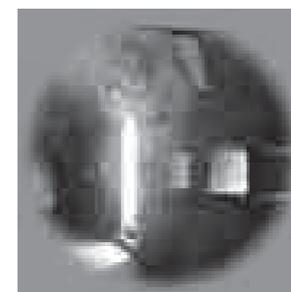
Faces



Fruits



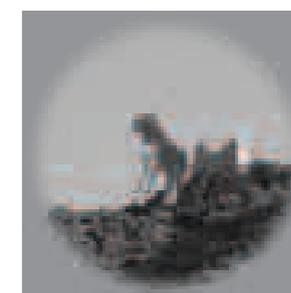
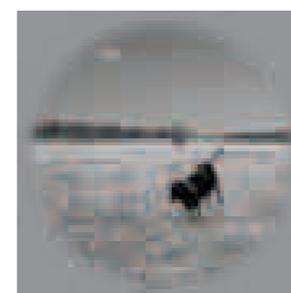
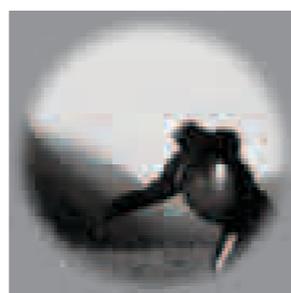
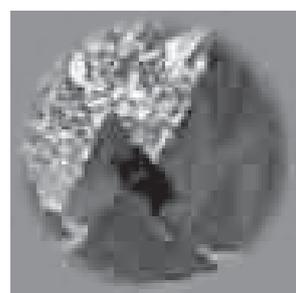
Planes



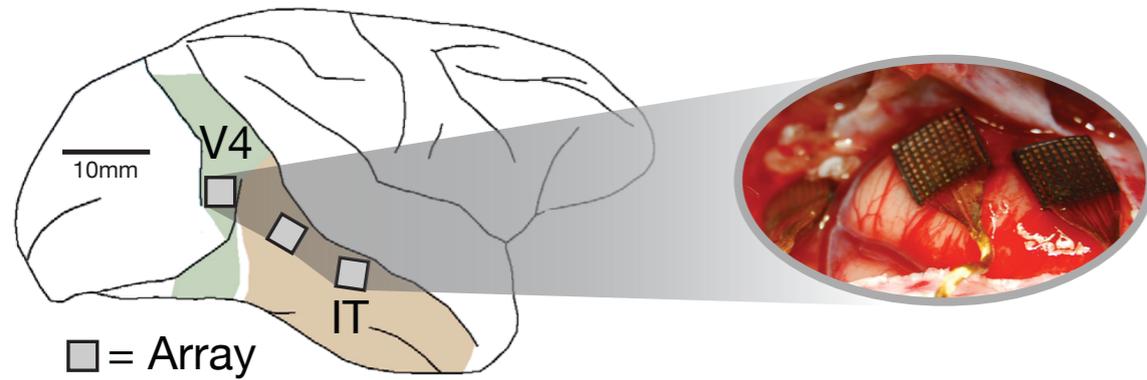
Tables



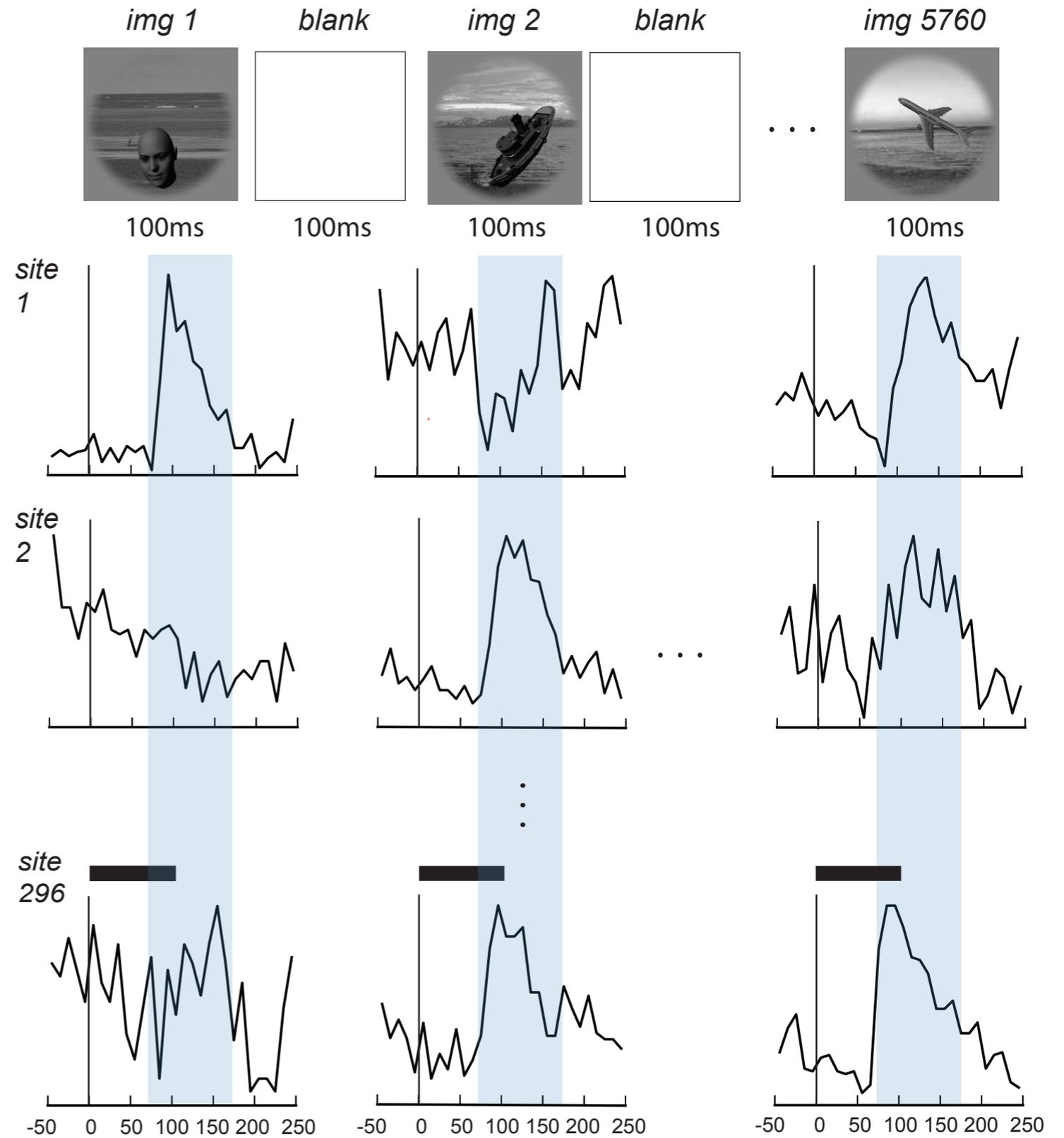
Pose, position, scale, and background variation



Example 2: The Ventral Visual Pathway



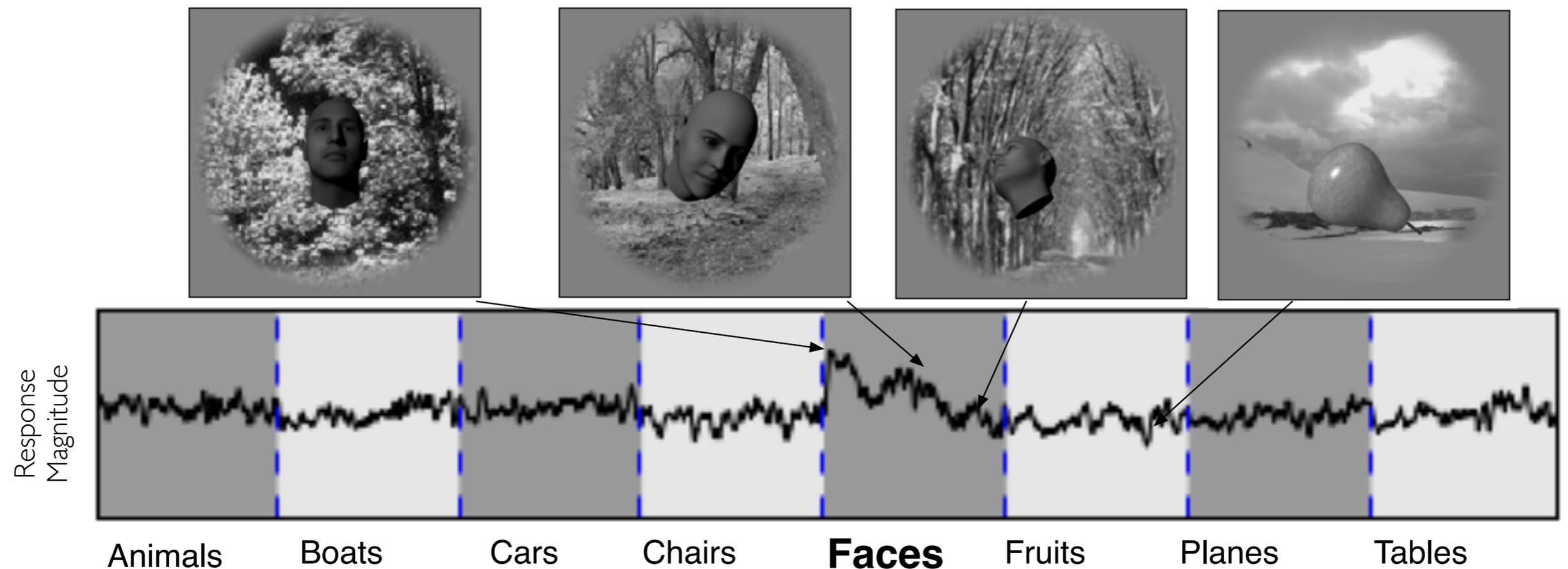
About 300 total sites



Example 2: The Ventral Visual Pathway

Responses to 1600 test images of two example units

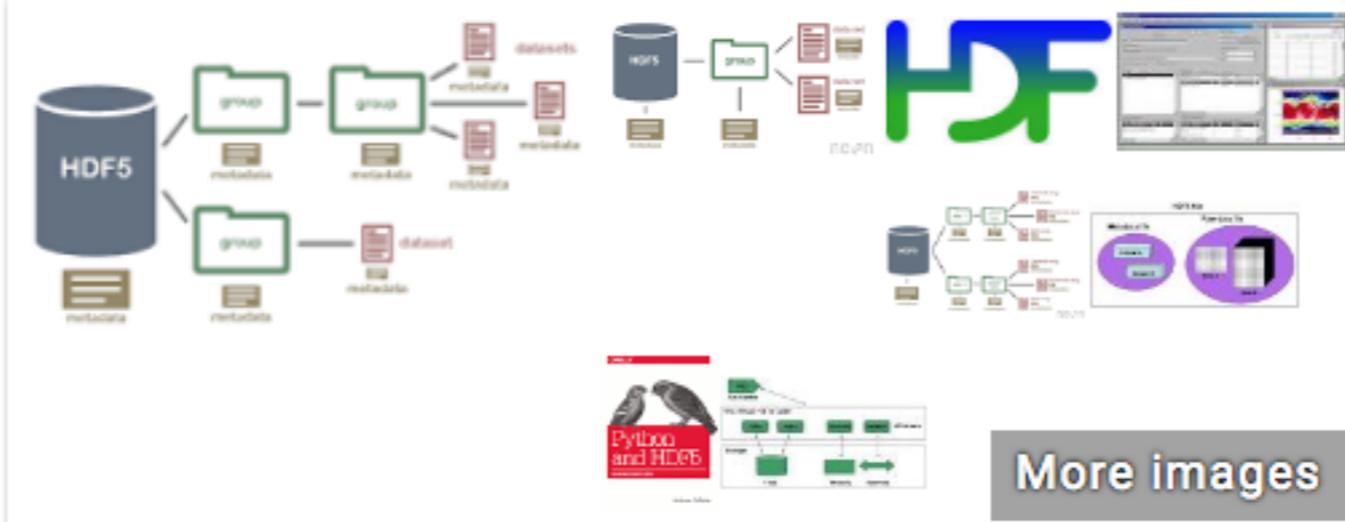
IT unit 53



Images sorted first by **category**, then **variation level**.

Example 2: The Ventral Visual Pathway

So what's the tensors in the data here?



Hierarchical Data Format

File format

Hierarchical Data Format is a set of file formats designed to store and organize large amounts of data. [Wikipedia](#)

Latest release: 5-1.10.1; (April 27, 2017; 10 months ago)

Magic number: \211HDF\r\n\032\n

Developed by: The HDF Group

HDF-5 data format

h5py python package

Example 2: The Ventral Visual Pathway

So what's the tensors in the data here?

<https://cran.r-project.org/web/packages/h5/vignettes/h5-Intro.html>

h5 - An Object Oriented Interface to HDF5

Mario Annau

2017-08-30

Introduction

The Hierarchical Data Format 5 (HDF5) is a binary data format and API created by the (HDF-Group 1997–2016) to better meet ever-increasing data storage demands of the scientific computing community. HDF5 files store homogeneous, multidimensional data sets organized in groups similar to the folder structure of a file system. As a self-describing file format HDF5 objects can be annotated with meta data using attributes. Compared to **R**'s integrated binary format HDF5 has various advantages.

- **Language Independence** HDF5 is implemented in C and includes APIs for a wide range of programming languages like e.g. C++, Fortran, *Python* and *Matlab*.
- **Partial I/O** HDF5 files support direct access to parts of the file without first parsing the entire contents, thus can process data sets not fitting into memory.}
- **Optimization** Access performance to parts of the HDF5 file can be further tuned by specifying the memory layout. The defined chunks can be cached in memory to further improve access times for subsequent queries.

Example 2: The Ventral Visual Pathway

So what's the tensors in the data here?

[IPYNB: Ventral stream neural data]

Example 2: The Ventral Visual Pathway

So what's the tensors in the data here?

1. The Stimuli (the images)

shape = (5760, 256, 256)
stimulus number rows cols (no color channels since gray-scale images)

2. The Stimulus Metadata (the images)

shape = (5760, # of attributes)

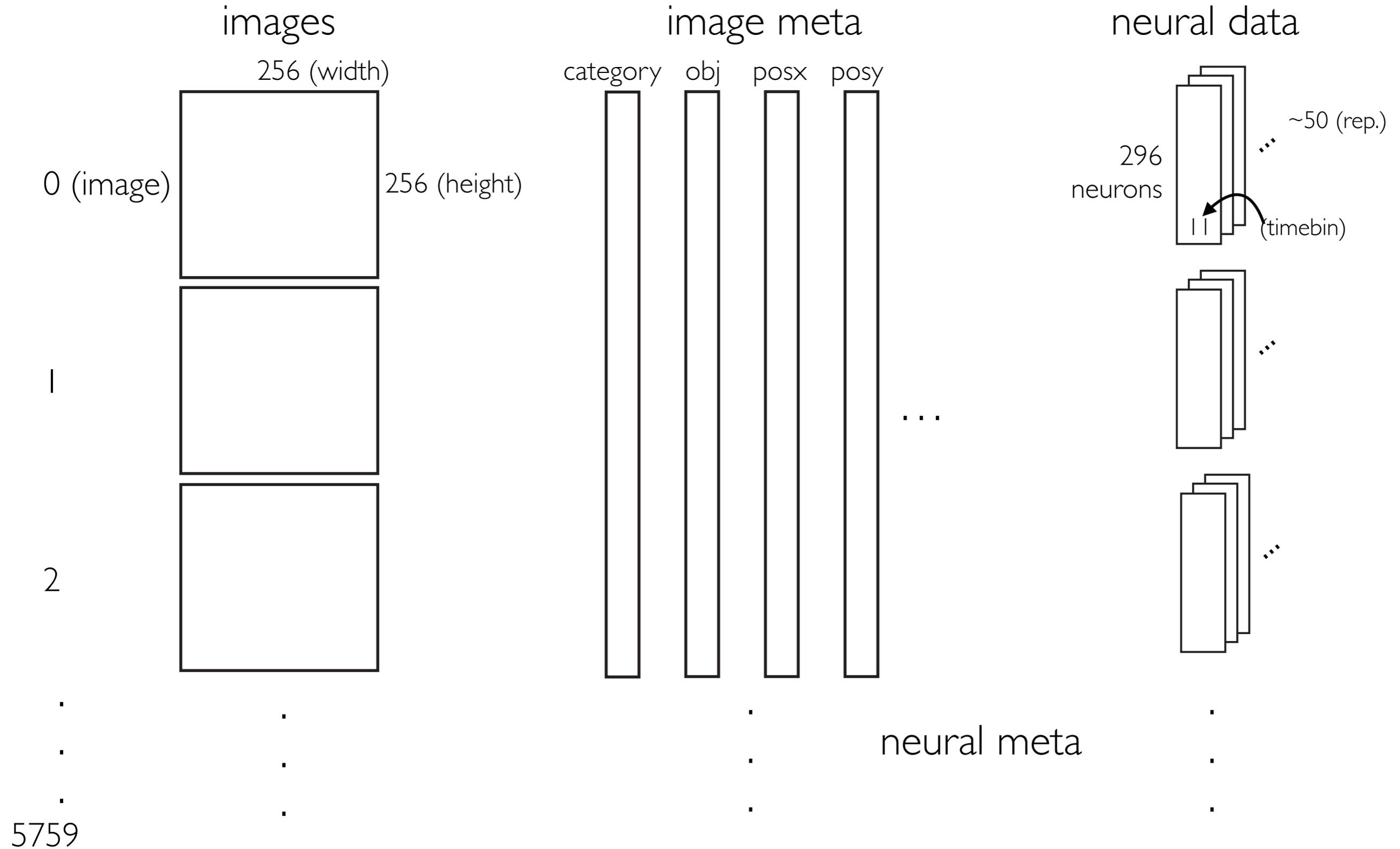
3. The "Raw" Neural Data

shape = (5760, 296, 10, ~50)
stimulus number neurons timebins repetitions of same image
(20 ms)
(not quite, since separated by var level)

+ various derived data tensors (time, trial averaging)

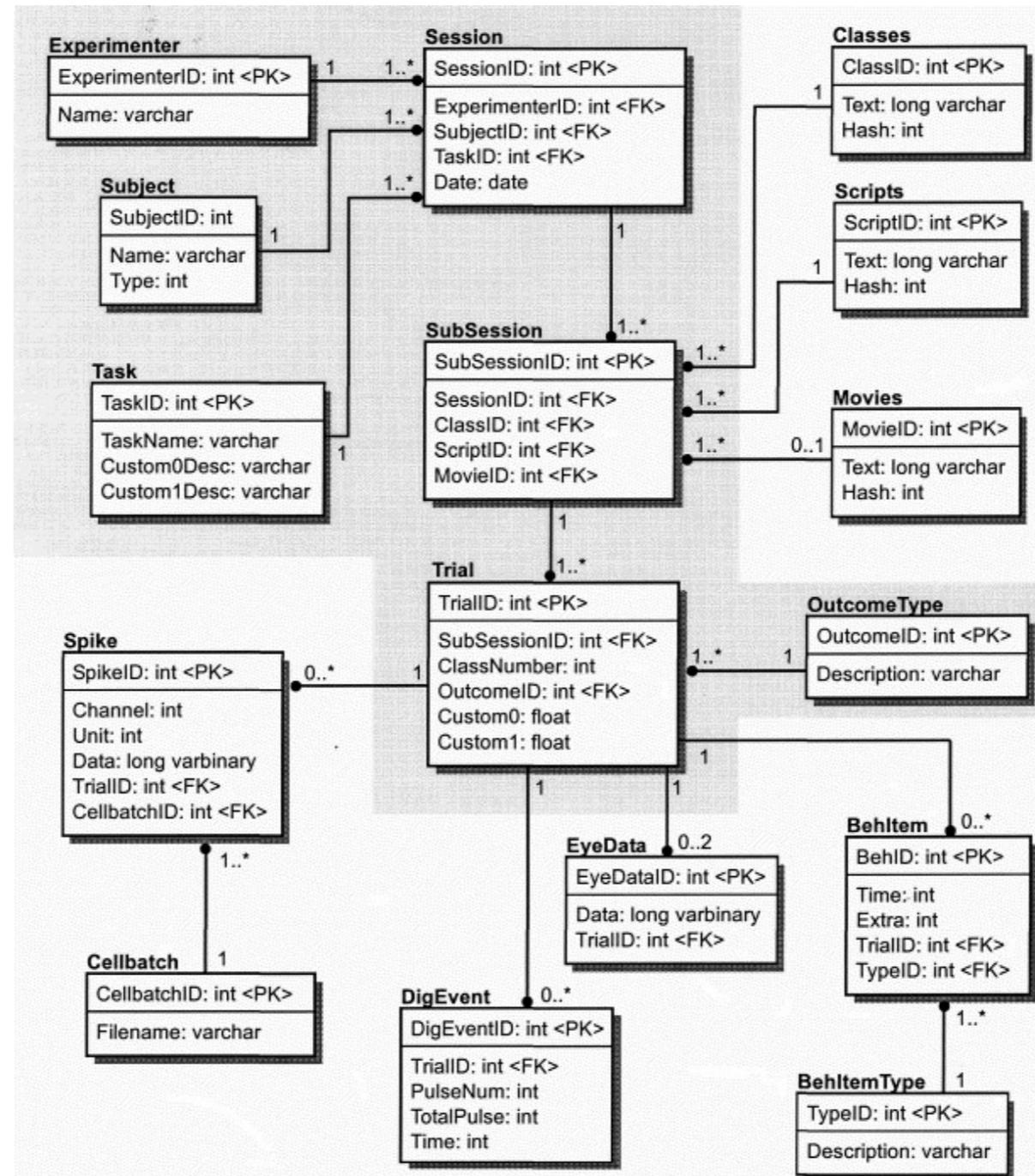
Example 2: The Ventral Visual Pathway

Names of dimensions informally create links between tensors.



Example 2: The Ventral Visual Pathway

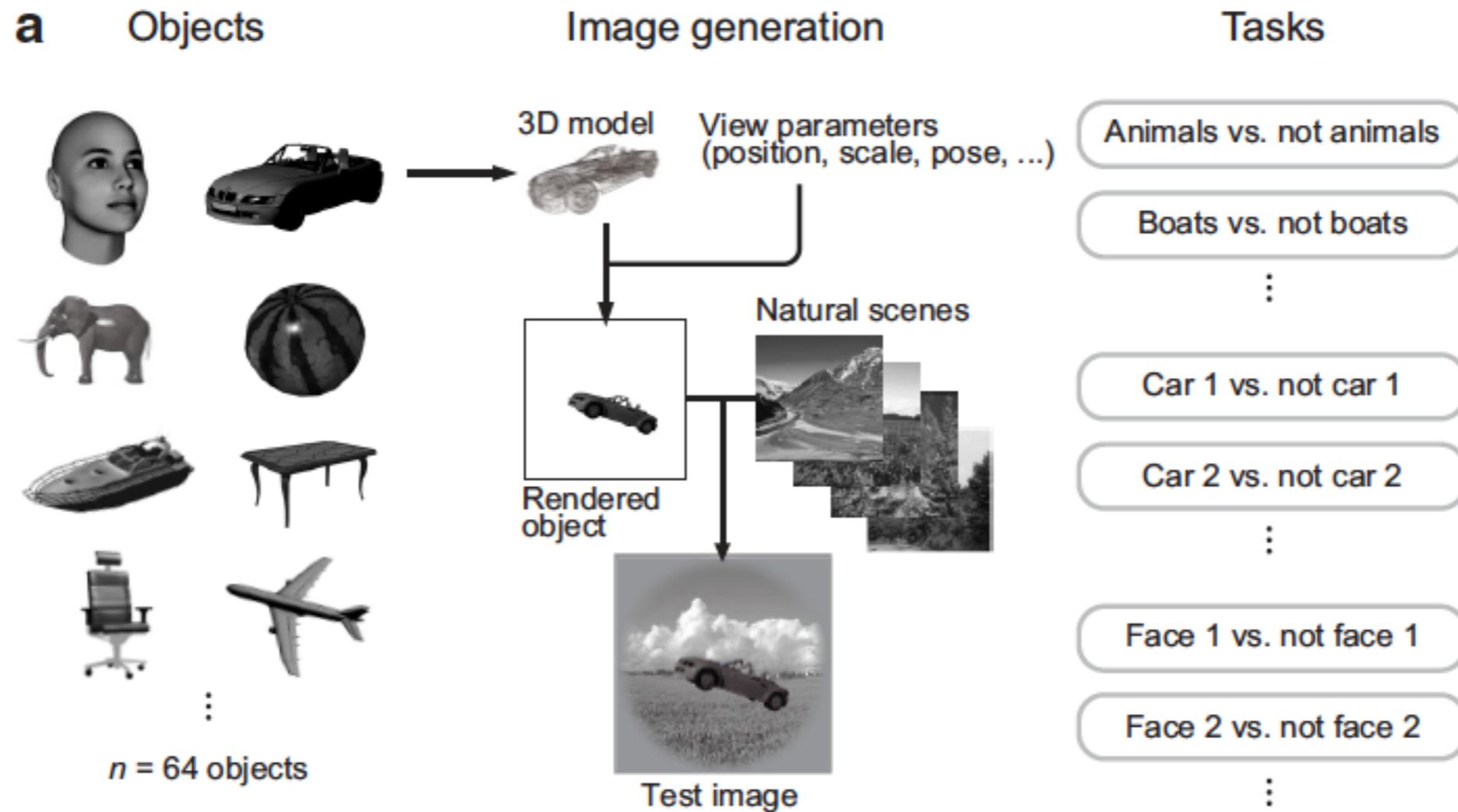
Relational databases are the “ultimate” linked object:



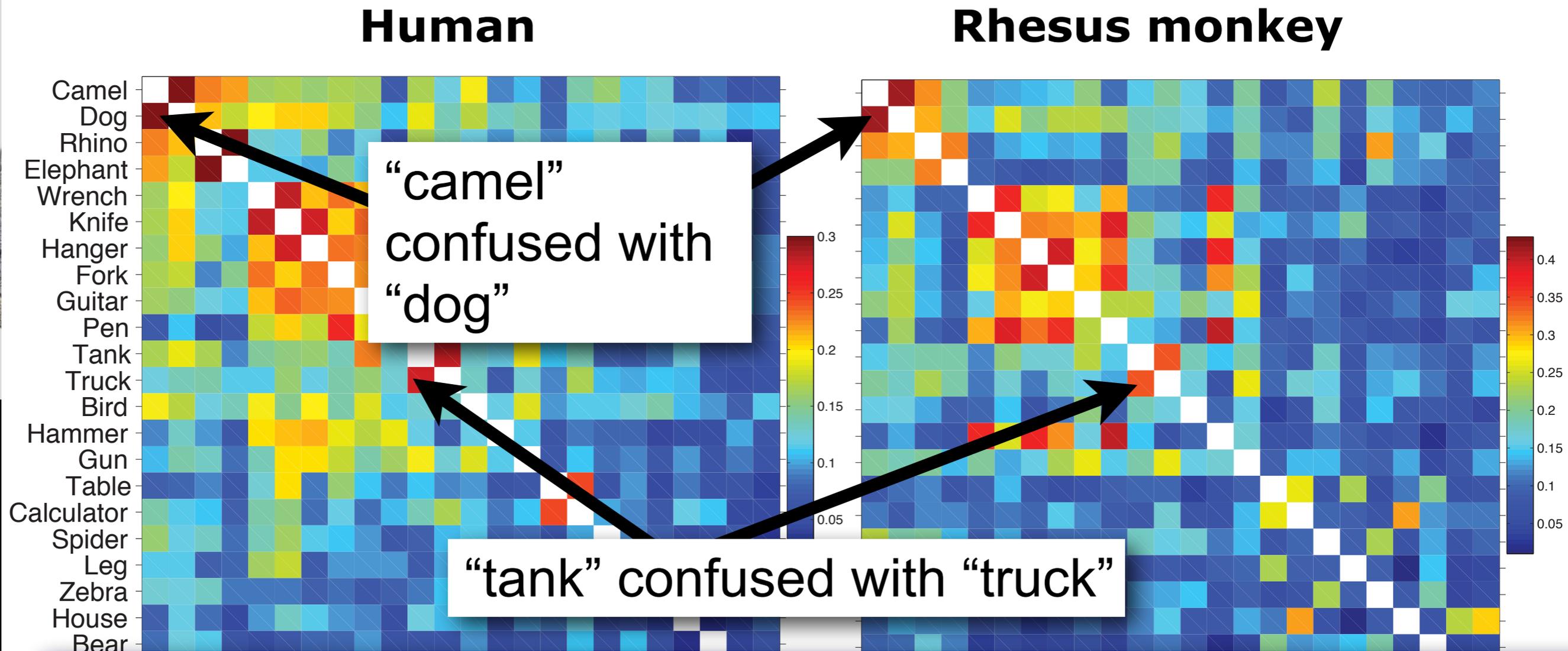
.... but often easier to work with “static” data representations like HDF5.

Example 2: The Ventral Visual Pathway

There's also direct behavioral data, in humans and monkeys.



Example 2: The Ventral Visual Pathway



Upshot: human and non-human primate basic level core object percept (sp. identification) are indistinguishable

Does not depend on reporting effector (touch vs. eye movement)

Comparison of Object Recognition Behavior in Human and Monkey

R. Rajalingham, K Schmidt, J.J. DiCarlo, **Vision Sciences Society** (2014)

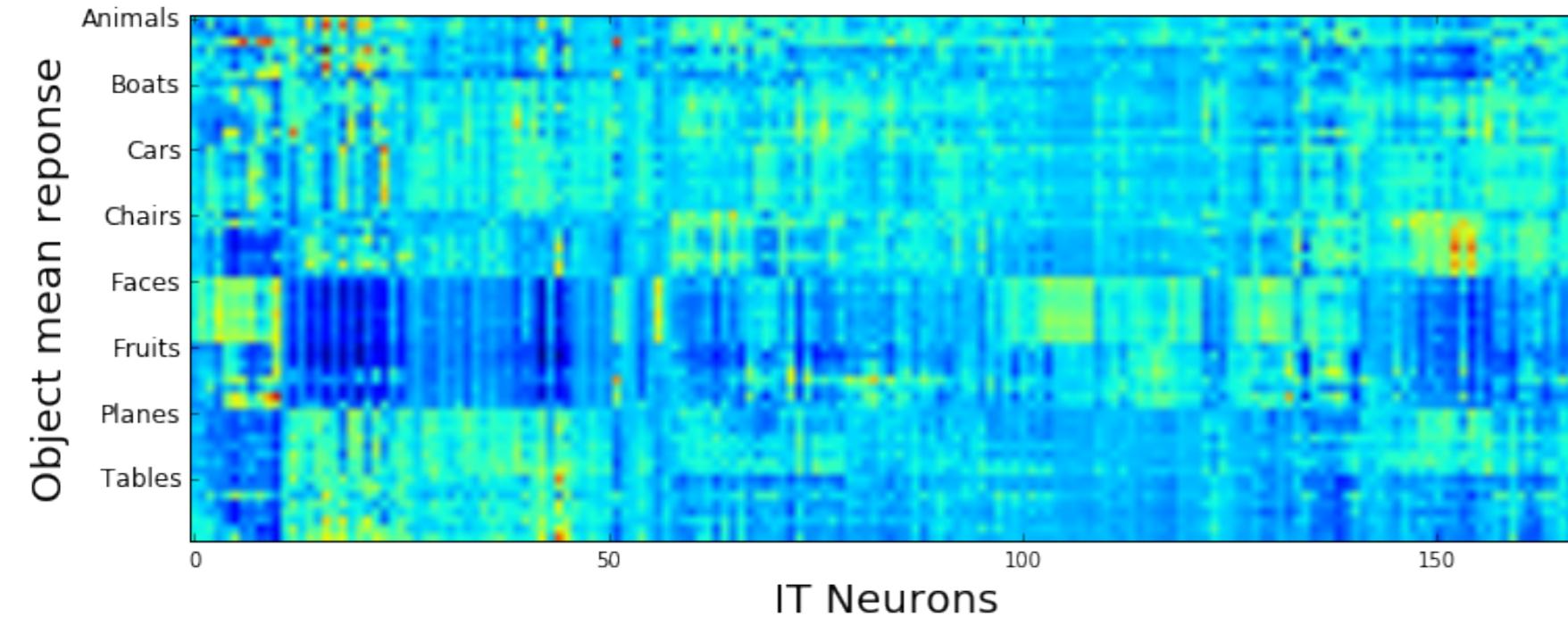
R. Rajalingham, K Schmidt, J.J. DiCarlo, **J. Neuroscience** (2015)

Example 2: The Ventral Visual Pathway

Examples of things we'll learn how to do:

IT neural representation

Behavioral patterns

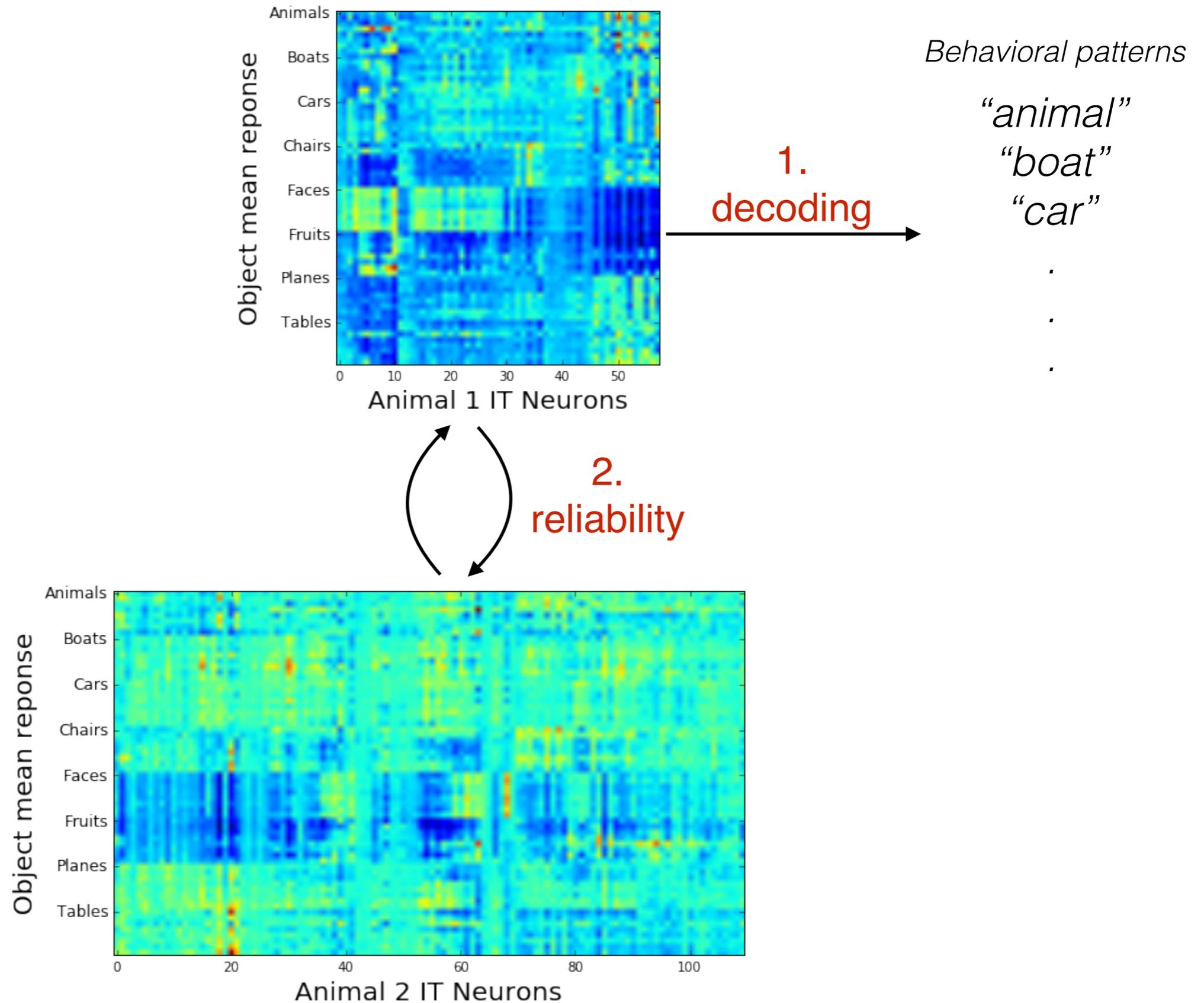


1.
decoding

“animal”
“boat”
“car”
.
.
.

Example 2: The Ventral Visual Pathway

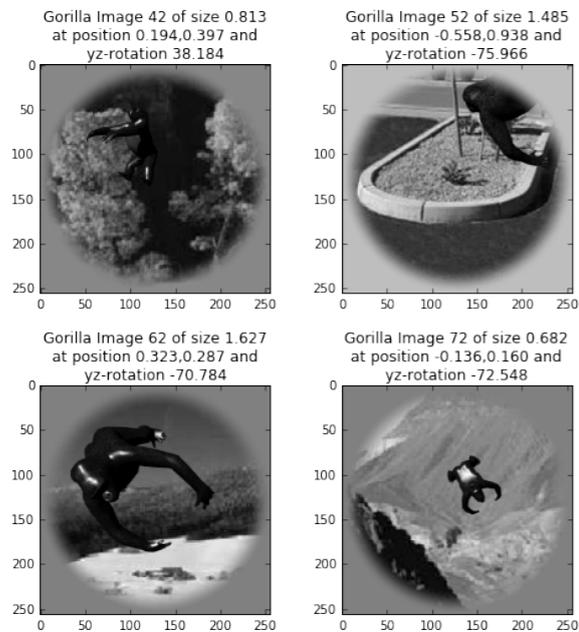
Examples of things we'll learn how to do:



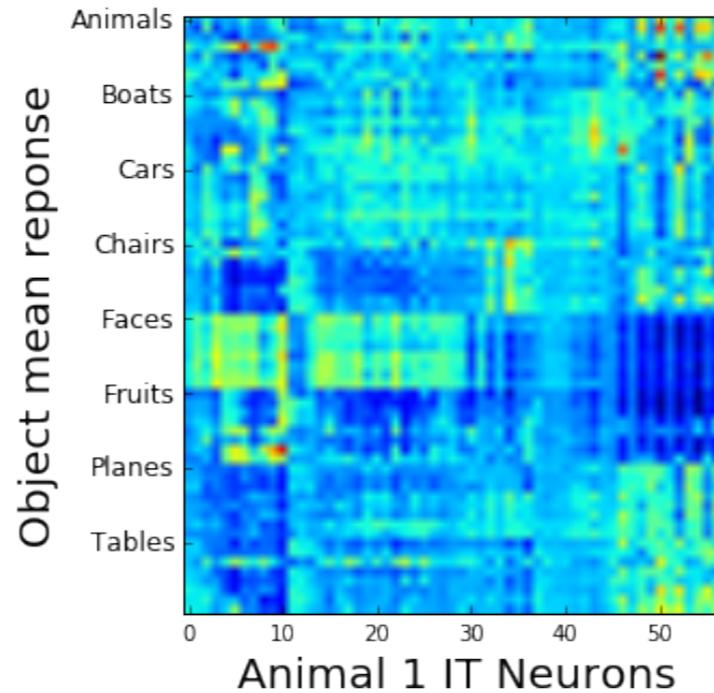
Example 2: The Ventral Visual Pathway

Examples of things we'll learn how to do:

The stimulus



3.
encoding

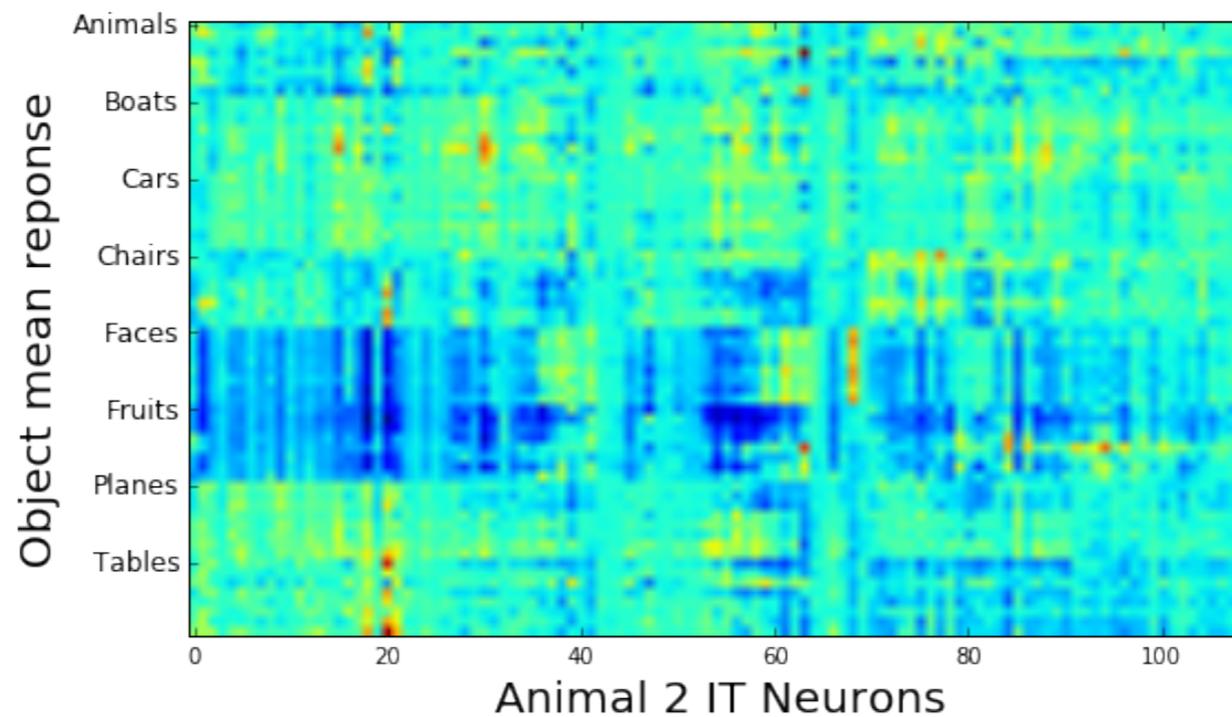


1.
decoding

Behavioral patterns

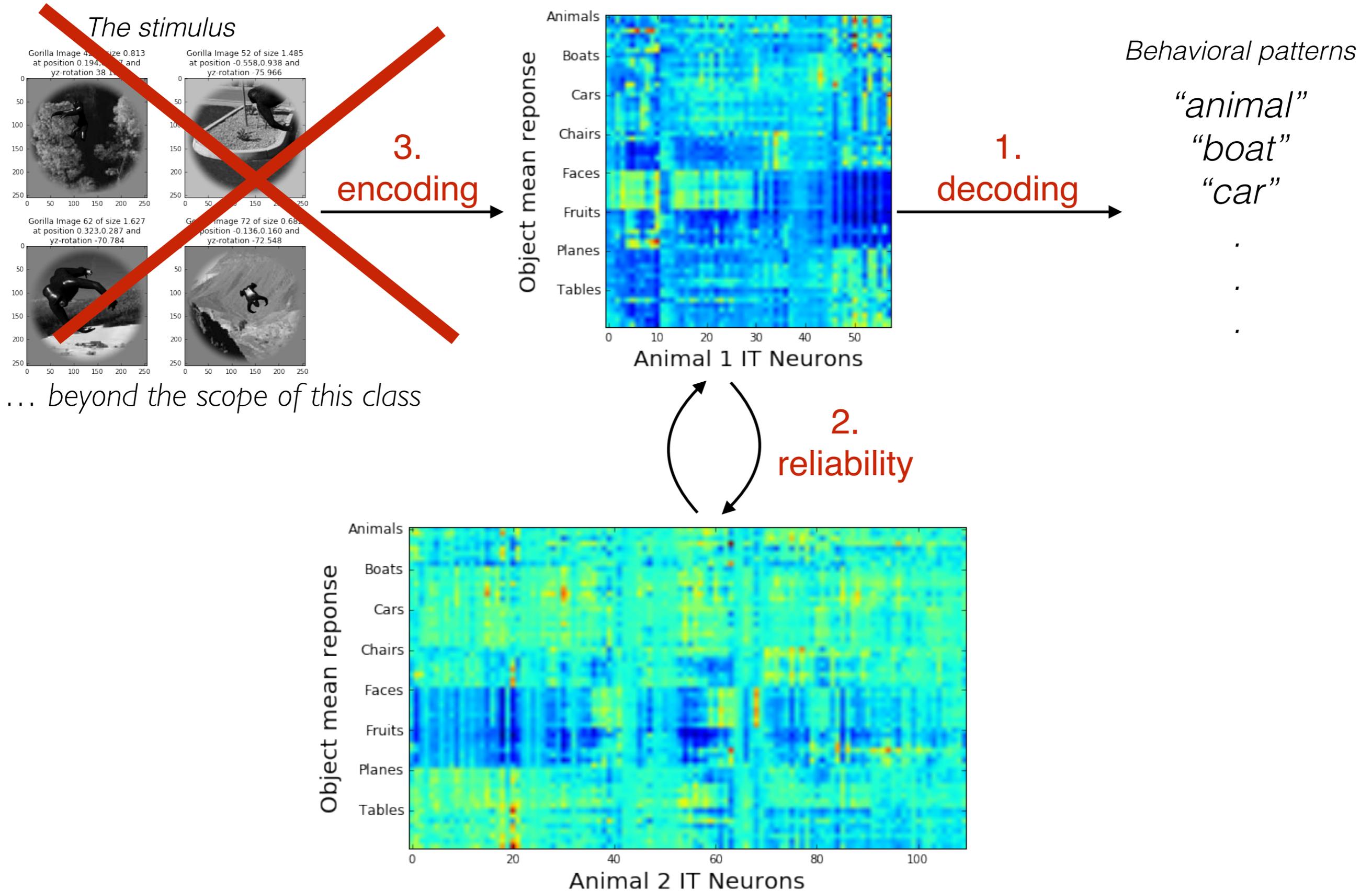
“animal”
“boat”
“car”
.
.
.

2.
reliability



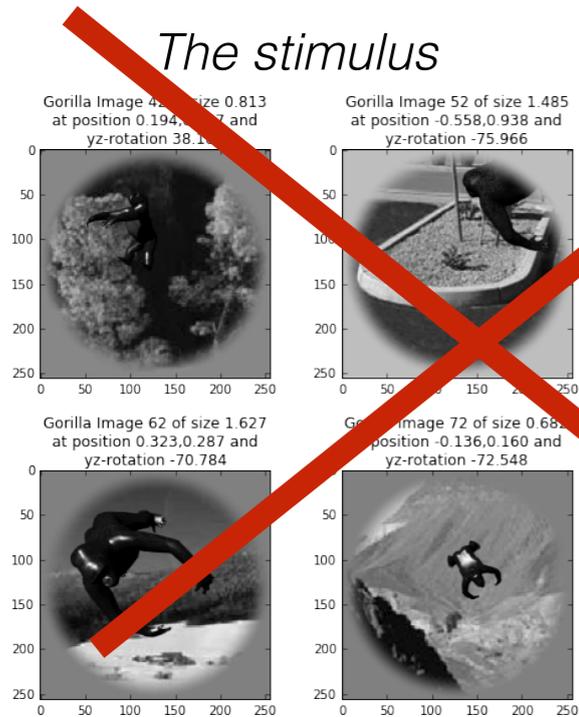
Example 2: The Ventral Visual Pathway

Examples of things we'll learn how to do:



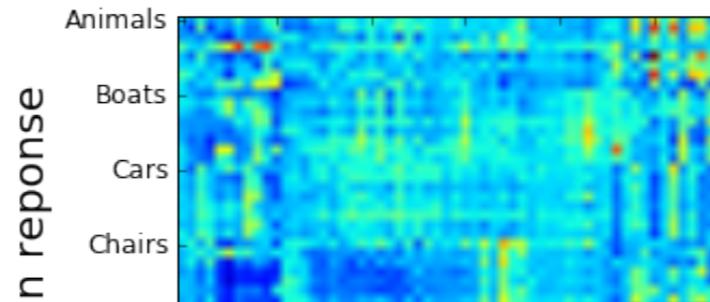
Example 2: The Ventral Visual Pathway

Examples of things we'll learn how to do:



... beyond the scope of this class

3. encoding

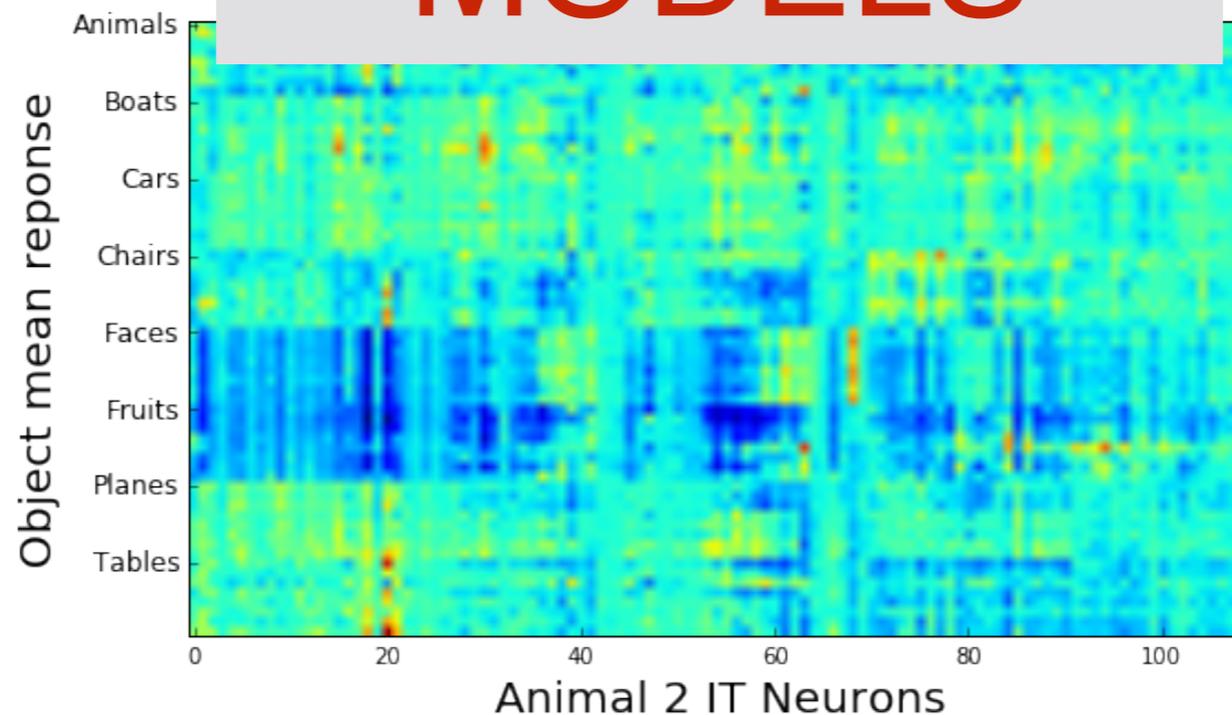


1. decoding

Behavioral patterns

“animal”
“boat”
“car”
.
.
.

ALL OF THESE THINGS ARE MODELS



What is a Model?

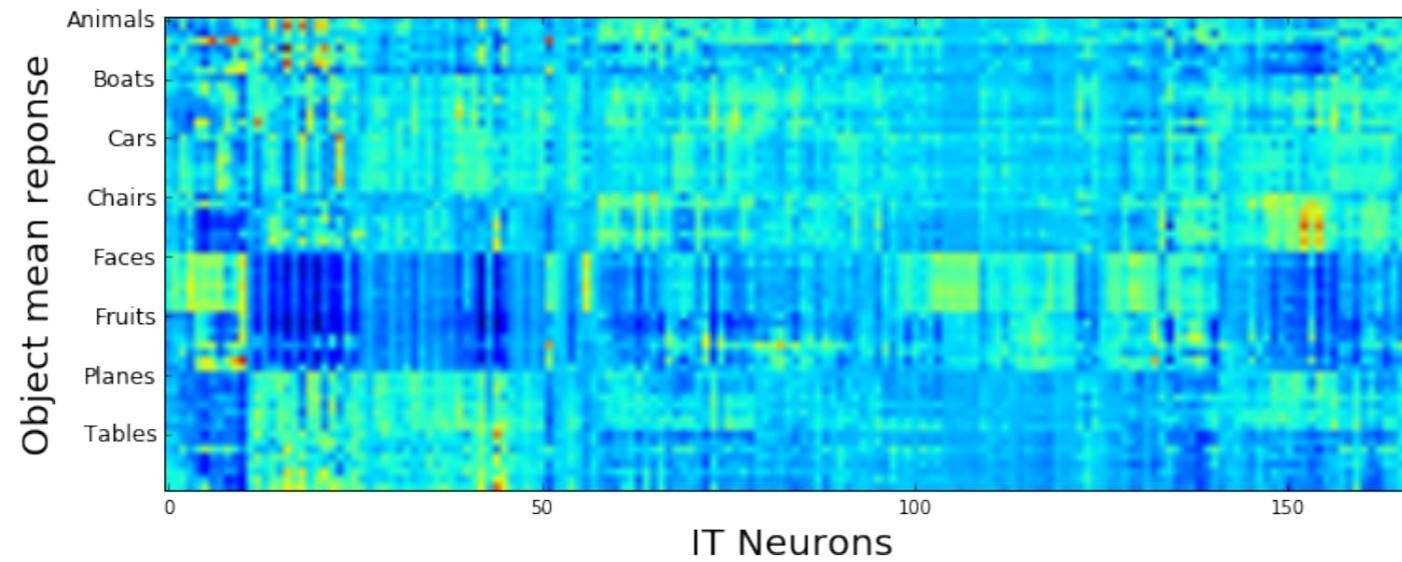
$$\text{Data Slice 2} = F(\text{data_slice_1})$$

$$\text{Data Slice 2} = F_{\text{params}}(\text{data_slice_1})$$

$$\text{Data Slice 2} = F_{\text{params}}(\text{data_slice_1}) + \text{Noise}$$

$$\text{Data Slice 2} = F_{\text{params}}(\text{data_slice_1})$$

IT neural representation



**1.
decoding**

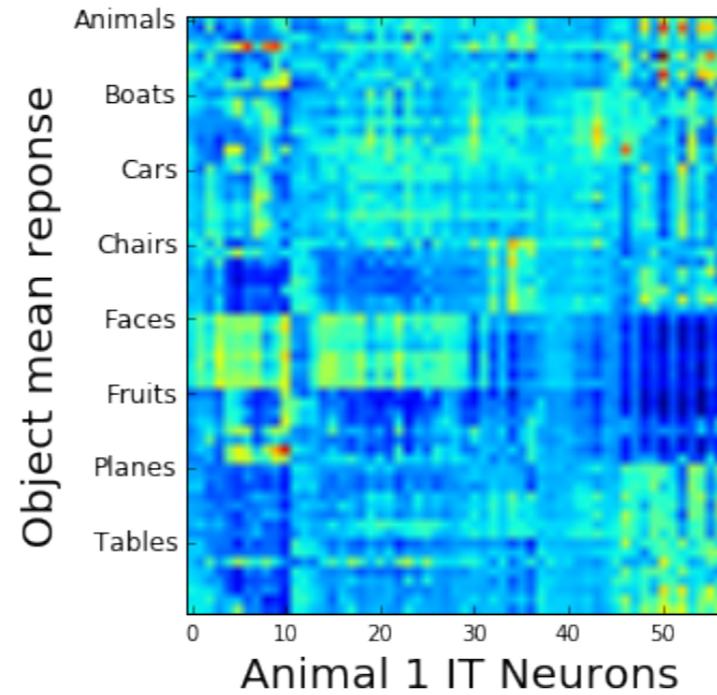
data_slice_1
=
neural data

Behavioral patterns

"animal"
"boat"
"car"

.
. .
data_slice_2
=
behavioral
data

$$\text{Data Slice 2} = F_{\text{params}}(\text{data_slice_1})$$

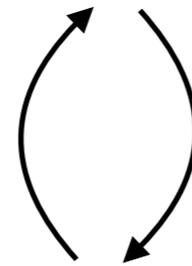


1.
decoding

Behavioral patterns

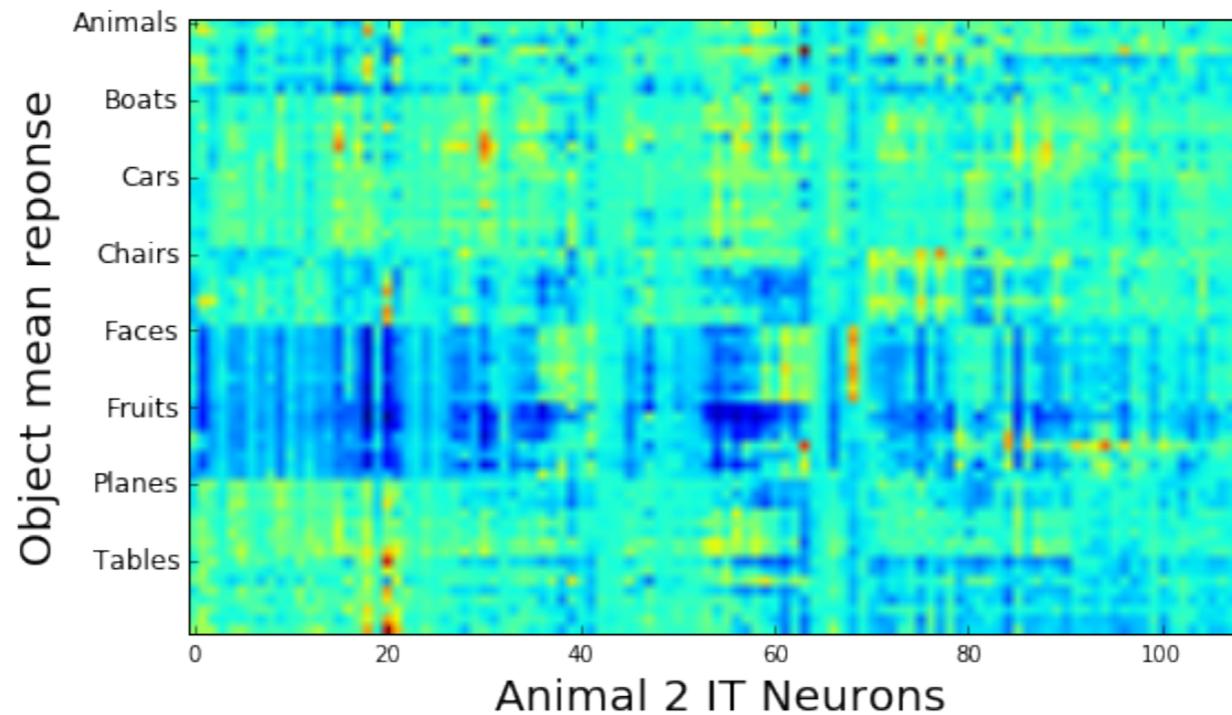
“animal”
“boat”
“car”
.
.
.

data_slice_1
=
neural data



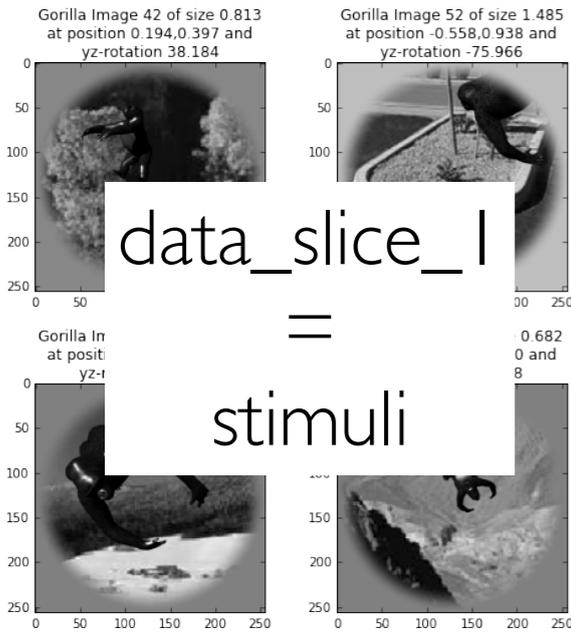
2.
reliability

data_slice_2
=
neural data

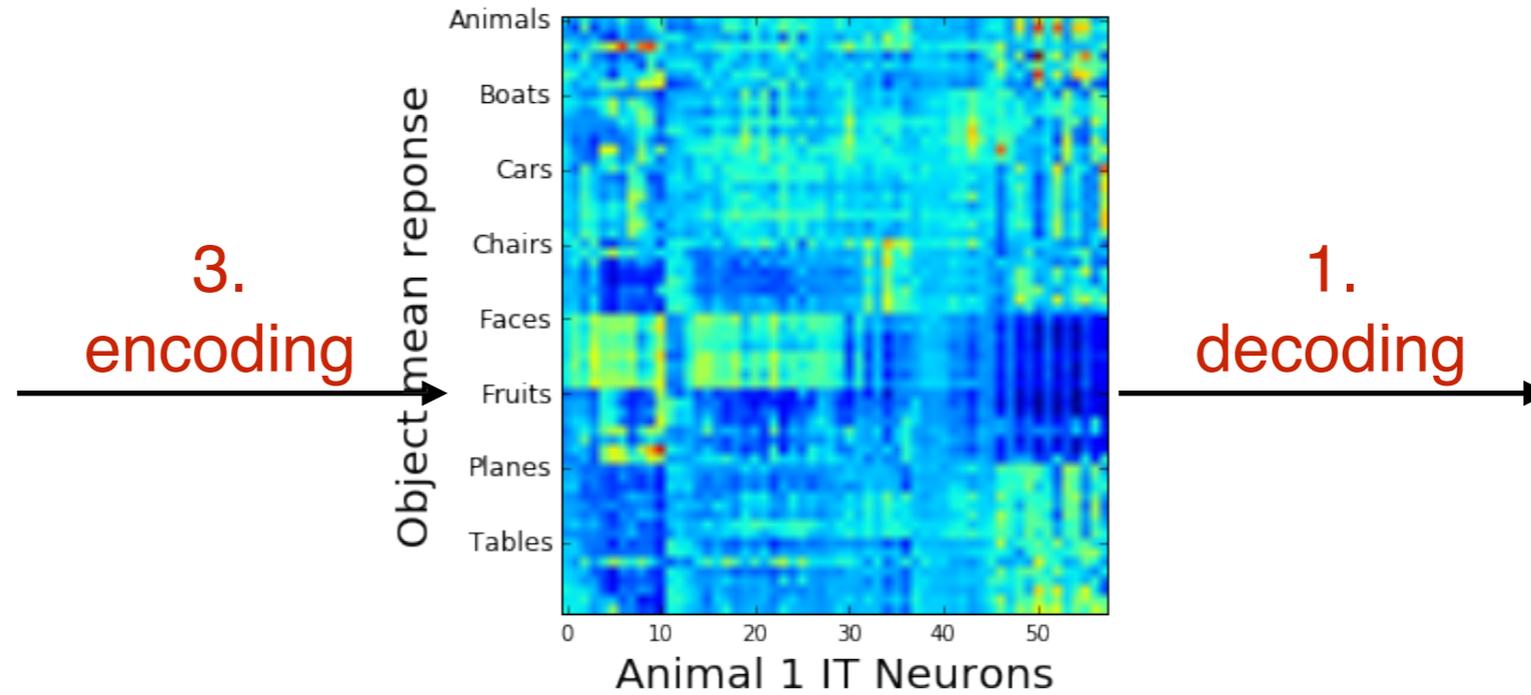


$$\text{Data Slice 2} = F_{\text{params}}(\text{data_slice_1})$$

The stimulus

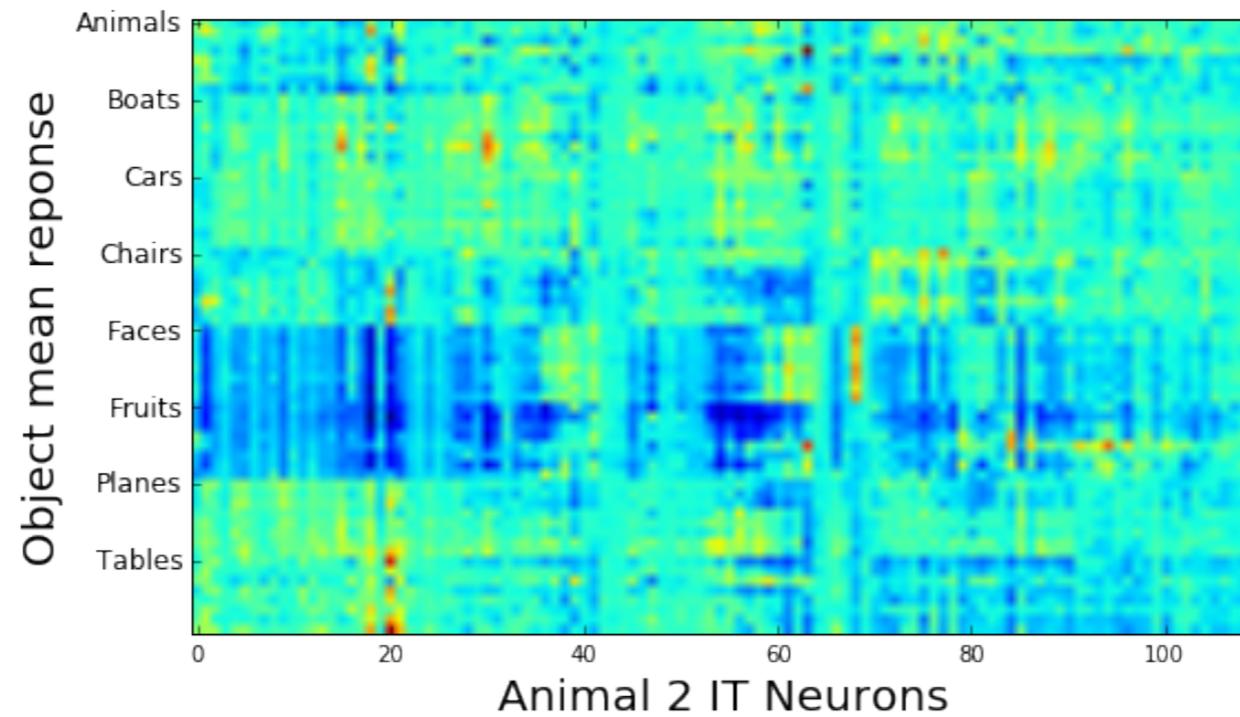
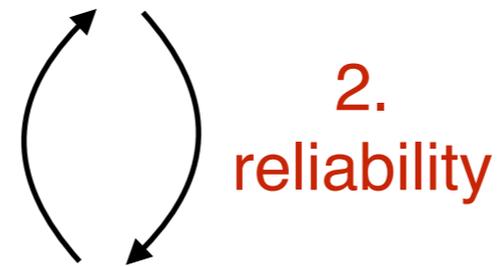


Neurons



Behavioral patterns

“animal”
“boat”
“car”
.
.
.



Example 3: The WordBank Dataset

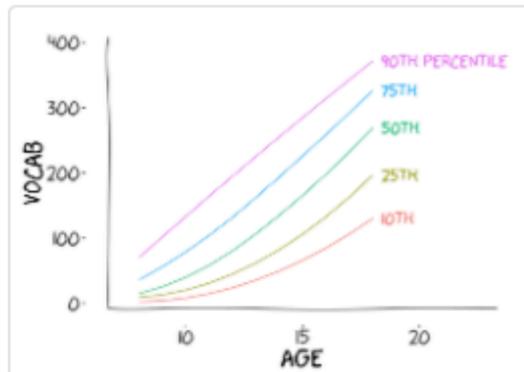


Wordbank

An open database of children's vocabulary development

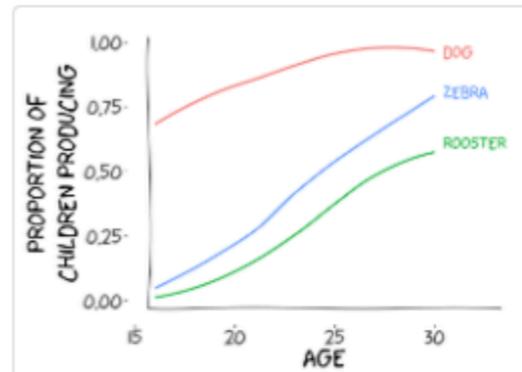


Wordbank contains data from 75,144 children and 82,983 CDI administrations, across 29 languages and 56 instruments:



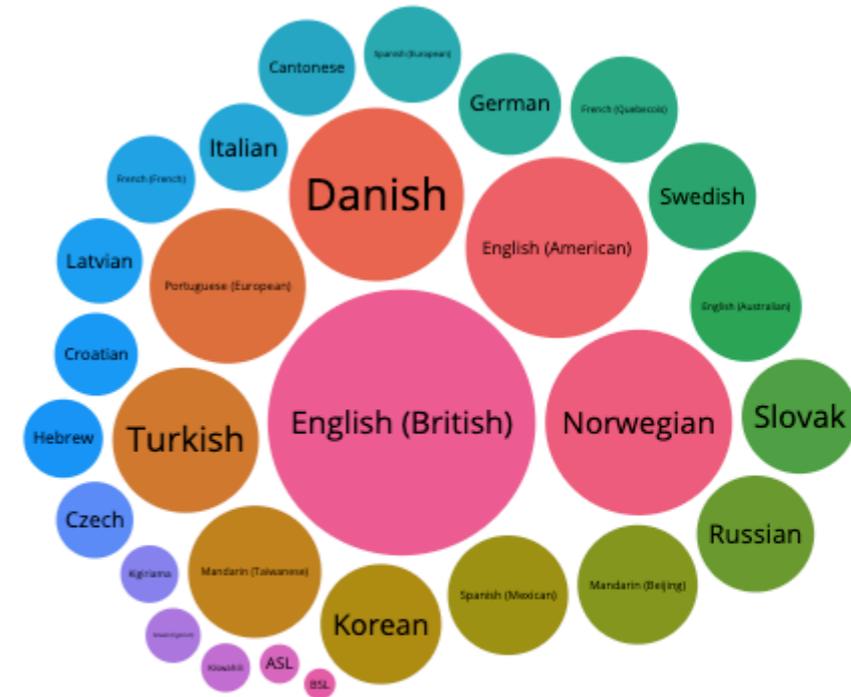
Vocabulary Norms

Explore vocabulary size growth curves for various languages and demographic groups.



Item Trajectories

Explore trajectories of individual words, word categories, and grammar items.



Wordbank is an open database of children's vocabulary growth, featuring data from [contributors around the world](#).

Example 3: The WordBank Dataset

Data Export Tools

By-Child Summary Data

data_id	language	form	age	sex	momed	comprehension	production
13141	English	WS	24	F	Graduate	337	337
13142	English	WS	19	F	College	384	384
13143	English	WS	24	M	Some Secondary	76	76
13144	English	WS	18	M	Secondary	19	19
13145	English	WS	24	F	Secondary	480	480
13146	English	WS	19	F	Some College	313	313
13147	English	WS	30	F	College	555	555
13148	English	WS	22	F	Some Secondary	505	505
13149	English	WS	18	M	Some Secondary	79	79
13150	English	WS	18	F	Secondary	346	346
13151	English	WS	18	F	Secondary	82	82
13152	English	WS	25	M	College	45	45
13153	English	WS	24	M	Some College	403	403

By-Word Summary Data

category	definition	16	17	18	19	20	21	22	23	24	25	
1	sounds	baa baa	0.47	0.57	0.61	0.68	0.71	0.62	0.78	0.77	0.79	0.82
2	sounds	choo choo	0.28	0.33	0.50	0.59	0.61	0.57	0.73	0.78	0.80	0.84
3	sounds	cockadoodleoo	0.13	0.11	0.19	0.25	0.27	0.33	0.39	0.38	0.50	0.53
4	sounds	grrr	0.54	0.57	0.63	0.66	0.66	0.63	0.75	0.80	0.83	0.79
5	sounds	meee	0.50	0.49	0.58	0.72	0.73	0.79	0.82	0.86	0.90	0.90
6	sounds	moo	0.55	0.61	0.71	0.76	0.78	0.76	0.85	0.91	0.90	0.90
7	sounds	ouch	0.43	0.45	0.60	0.65	0.73	0.75	0.75	0.87	0.84	0.87
8	sounds	quack quack	0.46	0.46	0.57	0.62	0.68	0.62	0.75	0.82	0.83	0.85
9	sounds	uh oh	0.76	0.78	0.87	0.90	0.90	0.89	0.90	0.93	0.94	0.94
10	sounds	vroom	0.49	0.46	0.59	0.63	0.57	0.65	0.68	0.72	0.76	0.73
11	sounds	woof woof	0.60	0.63	0.72	0.75	0.74	0.75	0.81	0.86	0.88	0.87
12	sounds	yum yum	0.56	0.53	0.61	0.64	0.64	0.58	0.66	0.73	0.72	0.77
13	animals	alligator	0.03	0.05	0.10	0.13	0.18	0.21	0.28	0.38	0.46	0.50
14	animals	animal	0.05	0.07	0.14	0.23	0.25	0.30	0.44	0.50	0.56	0.55
15	animals	ant	0.05	0.07	0.16	0.20	0.26	0.24	0.35	0.44	0.49	0.55
16	animals	bear	0.34	0.34	0.50	0.57	0.60	0.61	0.69	0.78	0.79	0.81
17	animals	bee	0.18	0.25	0.35	0.49	0.50	0.47	0.60	0.72	0.72	0.73
18	animals	bird	0.47	0.46	0.63	0.67	0.76	0.70	0.80	0.84	0.87	0.91

Full Child-by-Word Data

data_id	value	item_id	type	category	definition	age	sex	momed	comprehension	production
42962	produces	1	word	sounds	baa baa	24	F	Graduate	337	337
42962	produces	2	word	sounds	choo choo	24	F	Graduate	337	337
42962		3	word	sounds	cockadoodleoo	24	F	Graduate	337	337
42962	produces	4	word	sounds	grrr	24	F	Graduate	337	337
42962	produces	5	word	sounds	meow	24	F	Graduate	337	337
42962	produces	6	word	sounds	moo	24	F	Graduate	337	337
42962	produces	7	word	sounds	ouch	24	F	Graduate	337	337
42962	produces	8	word	sounds	quack quack	24	F	Graduate	337	337
42962	produces	9	word	sounds	uh oh	24	F	Graduate	337	337
42962		10	word	sounds	vroom	24	F	Graduate	337	337
42962	produces	11	word	sounds	woof woof	24	F	Graduate	337	337
42962	produces	12	word	sounds	yum yum	24	F	Graduate	337	337
42962		13	word	animals	alligator	24	F	Graduate	337	337
42962		14	word	animals	animal	24	F	Graduate	337	337
42962	produces	15	word	animals	ant	24	F	Graduate	337	337

Example 3: The WordBank Dataset

Data Export Tools

By-Child Summary Data

data_id	language	form	age	sex	momed	comprehension	production
13141	English	WS	24	F	Graduate	337	337
13142	English	WS	19	F	College	384	384
13143	English	WS	24	M	Some Secondary	76	76
13144	English	WS	18	M	Secondary	19	19
13145	English	WS	24	F	Secondary	480	480
13146	English	WS	19	F	Some College	313	313
13147	English	WS	30	F	College	555	555
13148	English	WS	22	F	Some Secondary	505	505
13149	English	WS	18	M	Some Secondary	79	79
13150	English	WS	18	F	Secondary	346	346
13151	English	WS	18	F	Secondary	82	82
13152	English	WS	25	M	College	45	45
13153	English	WS	24	M	Some College	403	403

By-Word Summary Data

category	definition	16	17	18	19	20	21	22	23	24	25	
1	sounds	baa baa	0.47	0.57	0.61	0.68	0.71	0.62	0.78	0.77	0.79	0.82
2	sounds	choo choo	0.28	0.33	0.50	0.59	0.61	0.57	0.73	0.78	0.80	0.84
3	sounds	cockadoodleoo	0.13	0.11	0.19	0.25	0.27	0.33	0.39	0.38	0.50	0.53
4	sounds	grrr	0.54	0.57	0.63	0.66	0.66	0.63	0.75	0.80	0.83	0.79
5	sounds	meee	0.50	0.49	0.58	0.72	0.73	0.79	0.82	0.86	0.90	0.90
6	sounds	moo	0.55	0.61	0.71	0.76	0.78	0.76	0.85	0.91	0.90	0.90
7	sounds	ouch	0.43	0.45	0.60	0.65	0.73	0.75	0.75	0.87	0.84	0.87
8	sounds	quack quack	0.46	0.46	0.57	0.62	0.68	0.62	0.75	0.82	0.83	0.85
9	sounds	uh oh	0.76	0.78	0.87	0.90	0.90	0.89	0.90	0.93	0.94	0.94
10	sounds	vroom	0.49	0.46	0.59	0.63	0.57	0.65	0.68	0.72	0.76	0.73
11	sounds	woof woof	0.60	0.63	0.72	0.75	0.74	0.75	0.81	0.86	0.88	0.87
12	sounds	yum yum	0.56	0.53	0.61	0.64	0.64	0.58	0.66	0.73	0.72	0.77
13	animals	alligator	0.03	0.05	0.10	0.13	0.18	0.21	0.28	0.38	0.46	0.50
14	animals	animal	0.05	0.07	0.14	0.23	0.25	0.30	0.44	0.50	0.56	0.55
15	animals	ant	0.05	0.07	0.16	0.20	0.26	0.24	0.35	0.44	0.49	0.55
16	animals	bear	0.34	0.34	0.50	0.57	0.60	0.61	0.69	0.78	0.79	0.81
17	animals	bee	0.18	0.25	0.35	0.49	0.50	0.47	0.60	0.72	0.72	0.73
18	animals	bird	0.47	0.46	0.63	0.67	0.76	0.70	0.80	0.84	0.87	0.91

Full Child-by-Word Data

data_id	value	item_id	type	category	definition	age	sex	momed	comprehension	production
42962	produces	1	word	sounds	baa baa	24	F	Graduate	337	337
42962	produces	2	word	sounds	choo choo	24	F	Graduate	337	337
42962		3	word	sounds	cockadoodleoo	24	F	Graduate	337	337
42962	produces	4	word	sounds	grrr	24	F	Graduate	337	337
42962	produces	5	word	sounds	meow	24	F	Graduate	337	337
42962	produces	6	word	sounds	moo	24	F	Graduate	337	337
42962	produces	7	word	sounds	ouch	24	F	Graduate	337	337
42962	produces	8	word	sounds	quack quack	24	F	Graduate	337	337
42962	produces	9	word	sounds	uh oh	24	F	Graduate	337	337
42962		10	word	sounds	vroom	24	F	Graduate	337	337
42962	produces	11	word	sounds	woof woof	24	F	Graduate	337	337
42962	produces	12	word	sounds	yum yum	24	F	Graduate	337	337
42962		13	word	animals	alligator	24	F	Graduate	337	337
42962		14	word	animals	animal	24	F	Graduate	337	337
42962	produces	15	word	animals	ant	24	F	Graduate	337	337

Static Reports

Gender Differences in Vocabulary

Maternal Education Differences in Vocabulary

Relationship between Lexicon and Grammar

Vocabulary Composition Trajectories

Estimating Words' Age of Acquisition

Predicting Words' Age of Acquisition

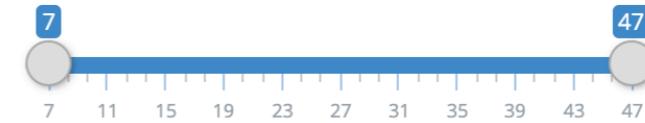
Example 3: The WordBank Dataset

By-Child Summary Data

Language:

Form:

Age (Months):



Gender:

Maternal Education:

[Download Data](#)

Show entries

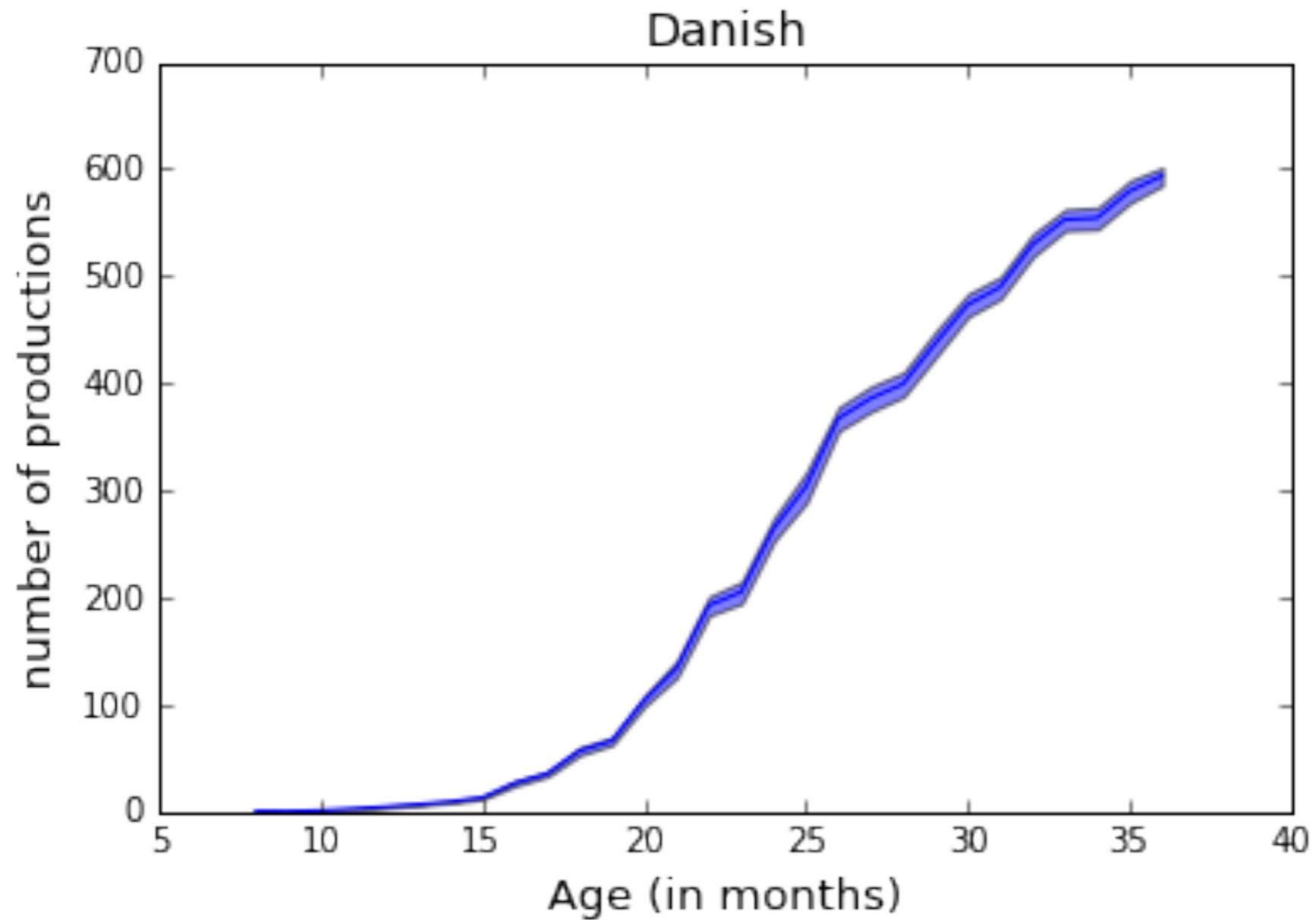
Search:

data_id	language	form	age	gender	mom_ed	comprehension	production
29821	Croatian	WG	13	Female		293	88
29822	Croatian	WG	16	Male		122	12
29823	Croatian	WG	9	Female		3	0
29824	Croatian	WG	12	Female		0	0
29825	Croatian	WG	12	Female		44	0
29826	Croatian	WG	8	Male		14	5
29827	Croatian	WG	9	Male		2	1
29828	Croatian	WG	10	Male		44	1
29829	Croatian	WG	13	Male		172	51
29830	Croatian	WG	16	Male		241	68
29831	Croatian	WG	16	Male		103	14
29832	Croatian	WG	16	Female		76	10

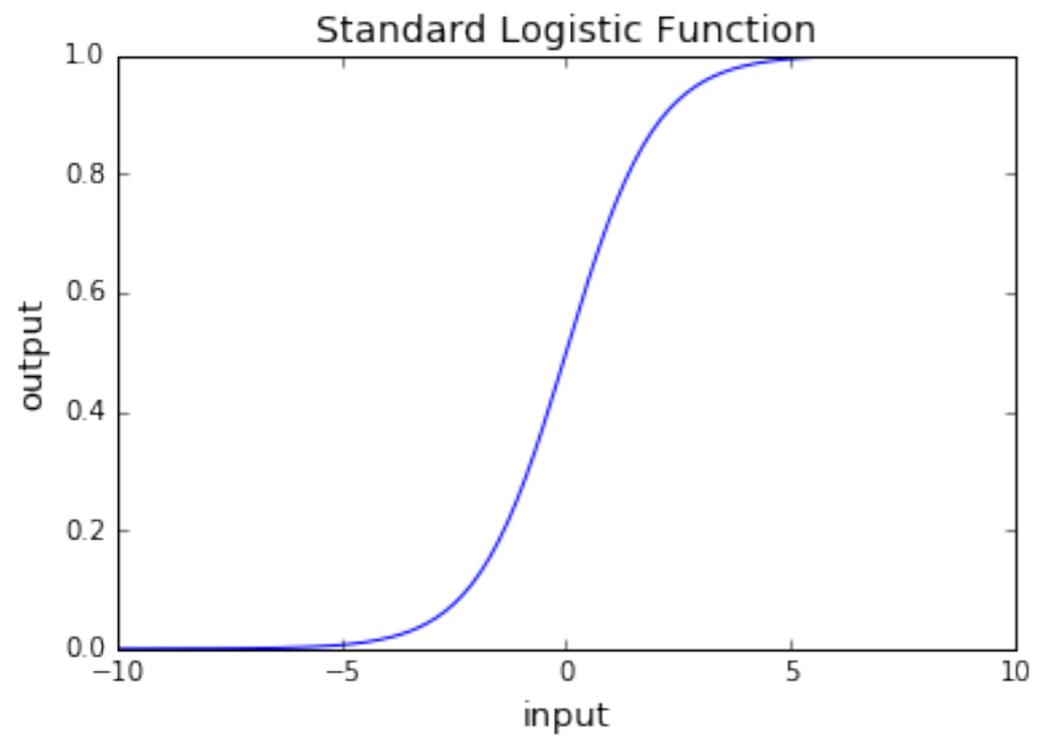
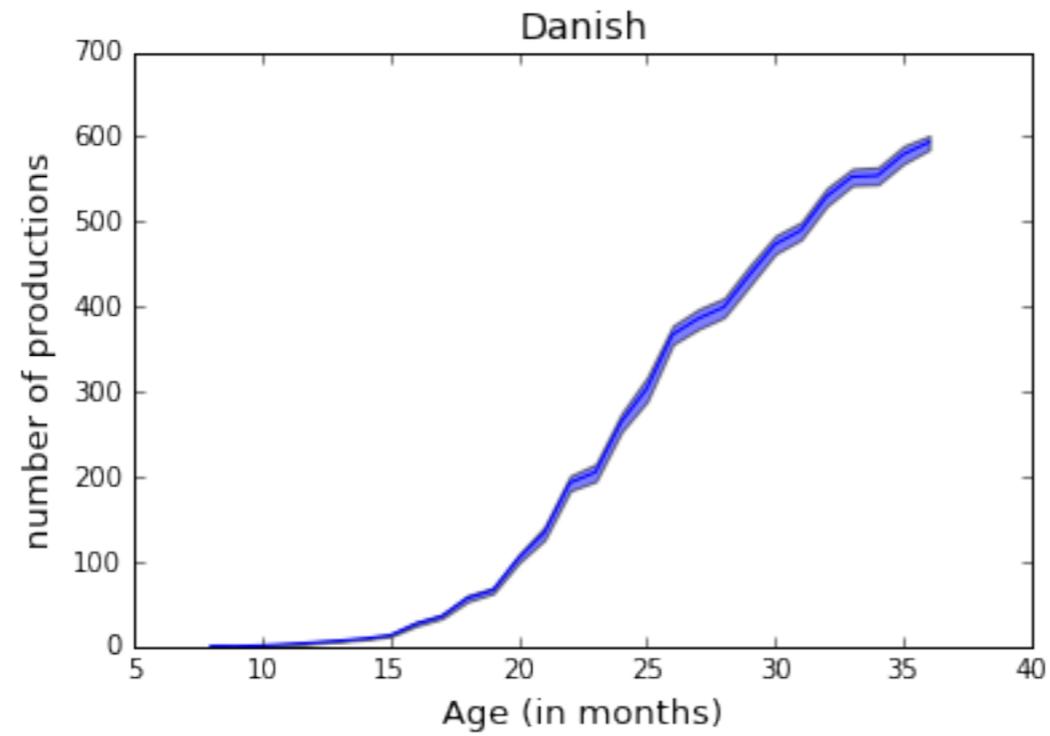
Example 3: The WordBank Dataset

[IPYNB: the Wordbank dataset]

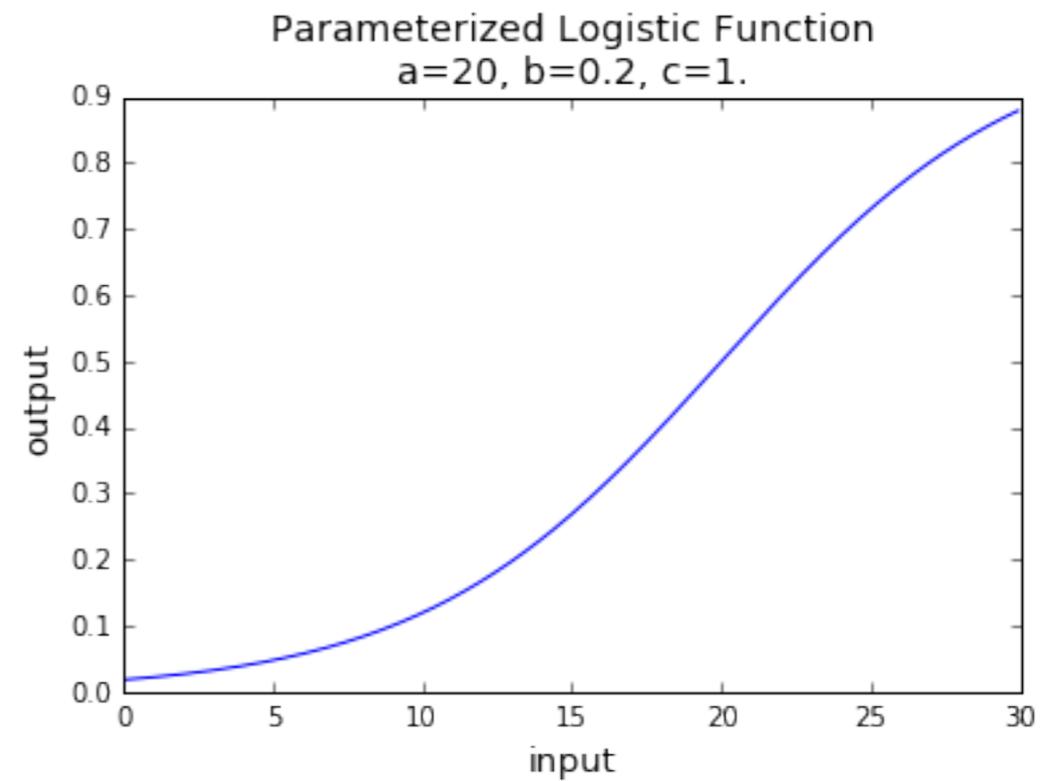
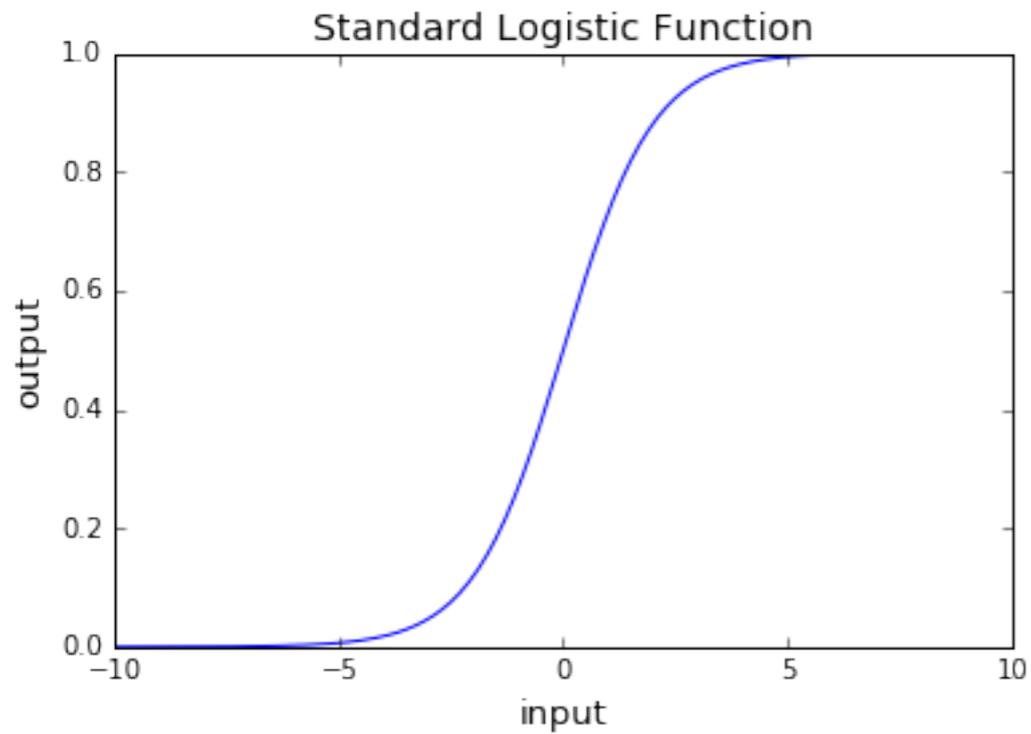
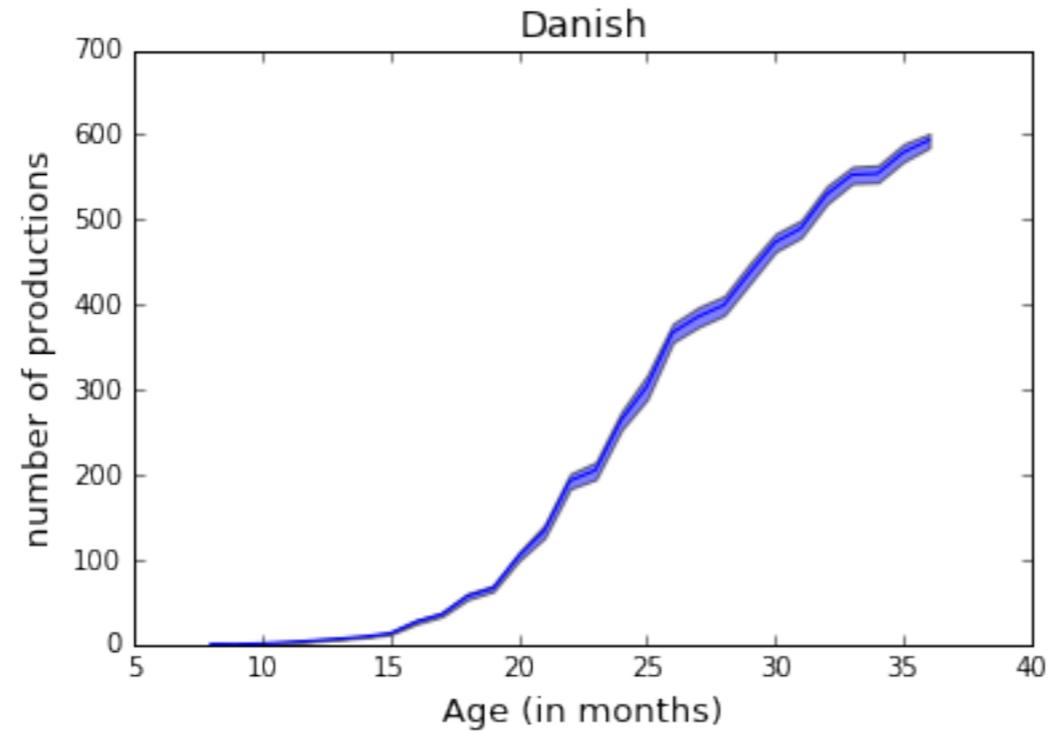
Example 3: The WordBank Dataset



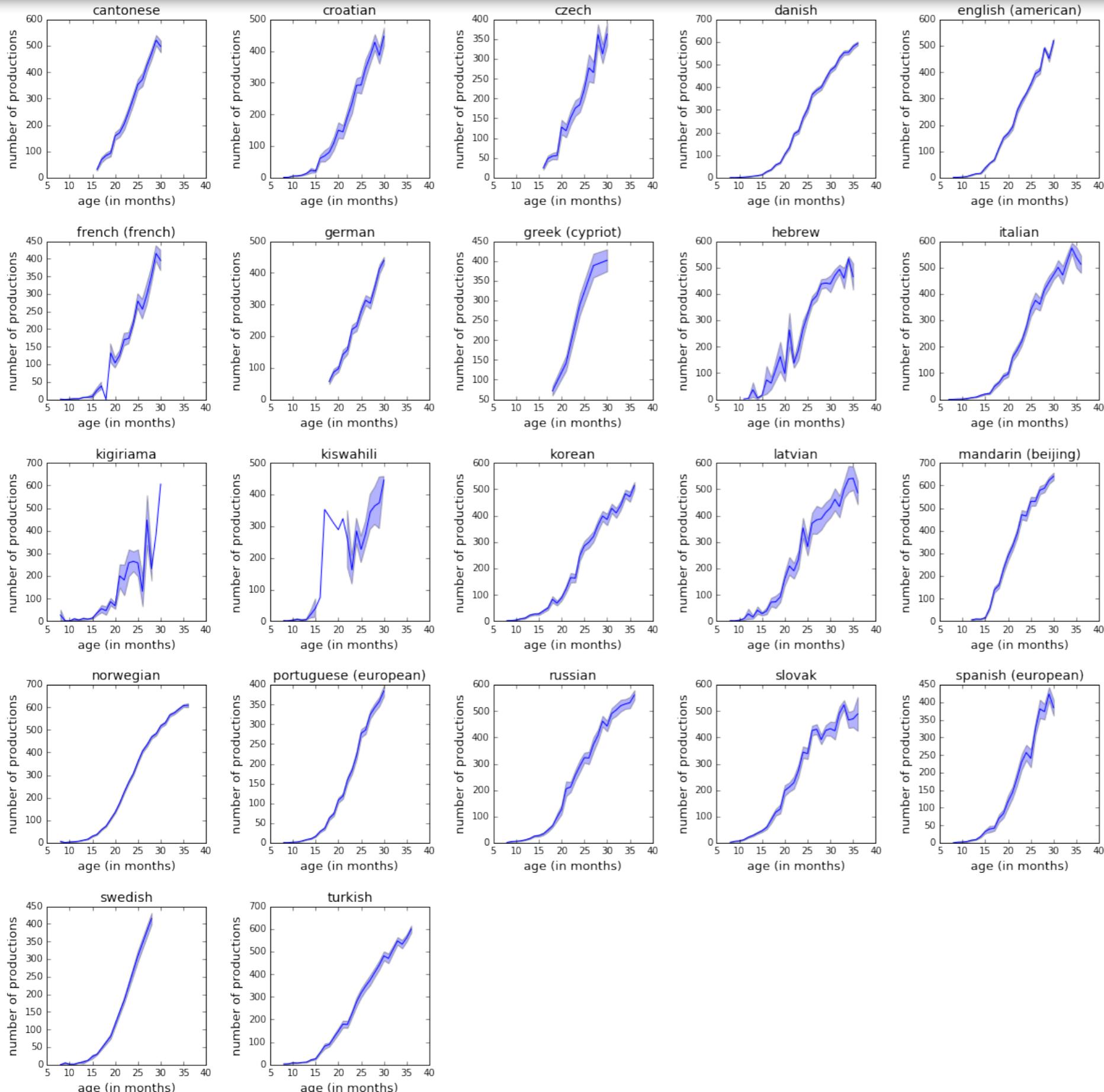
Example 3: The WordBank Dataset



Example 3: The WordBank Dataset

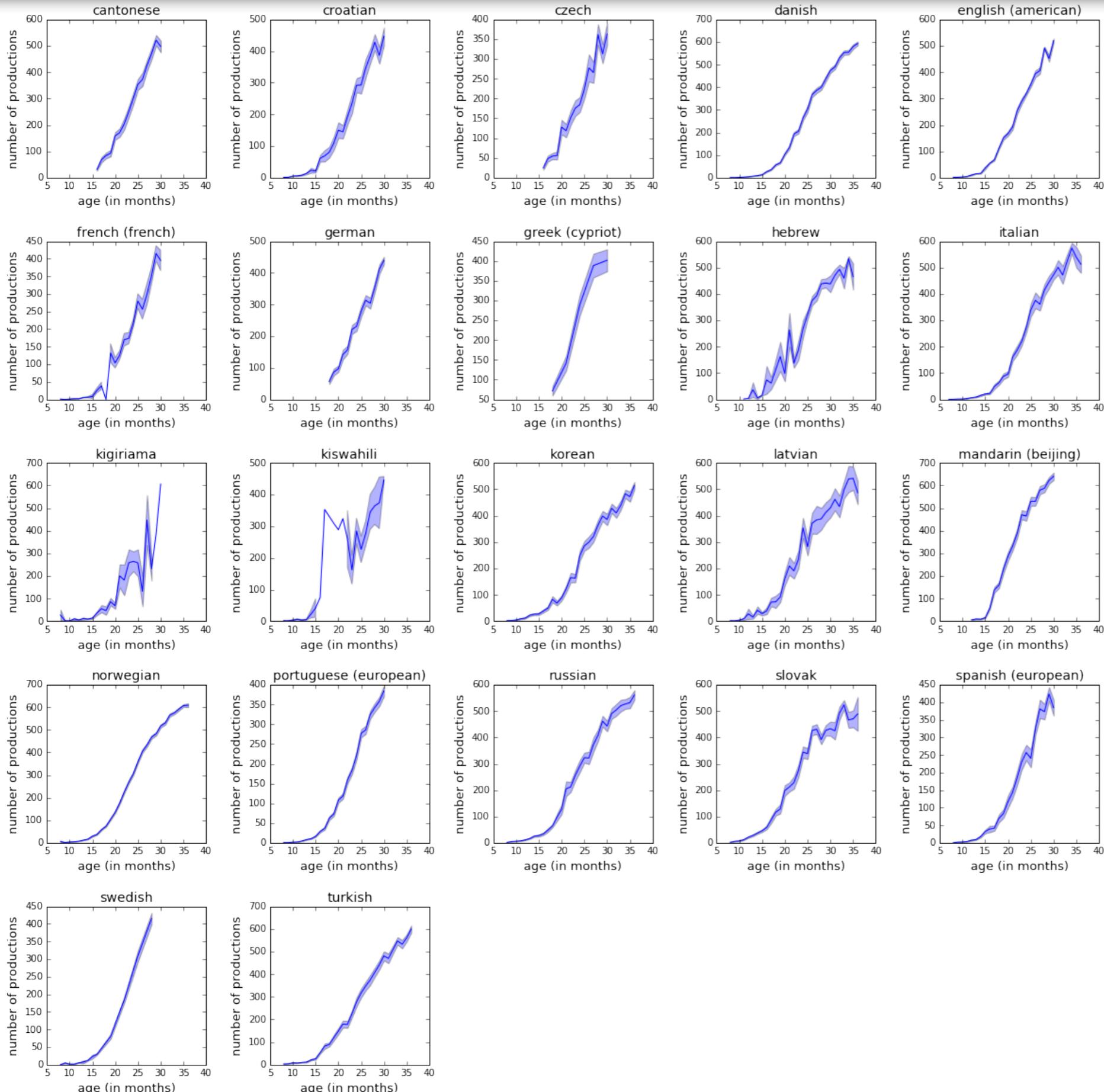


Example 3: The WordBank Dataset



Is the functional form the same across languages?

Example 3: The WordBank Dataset



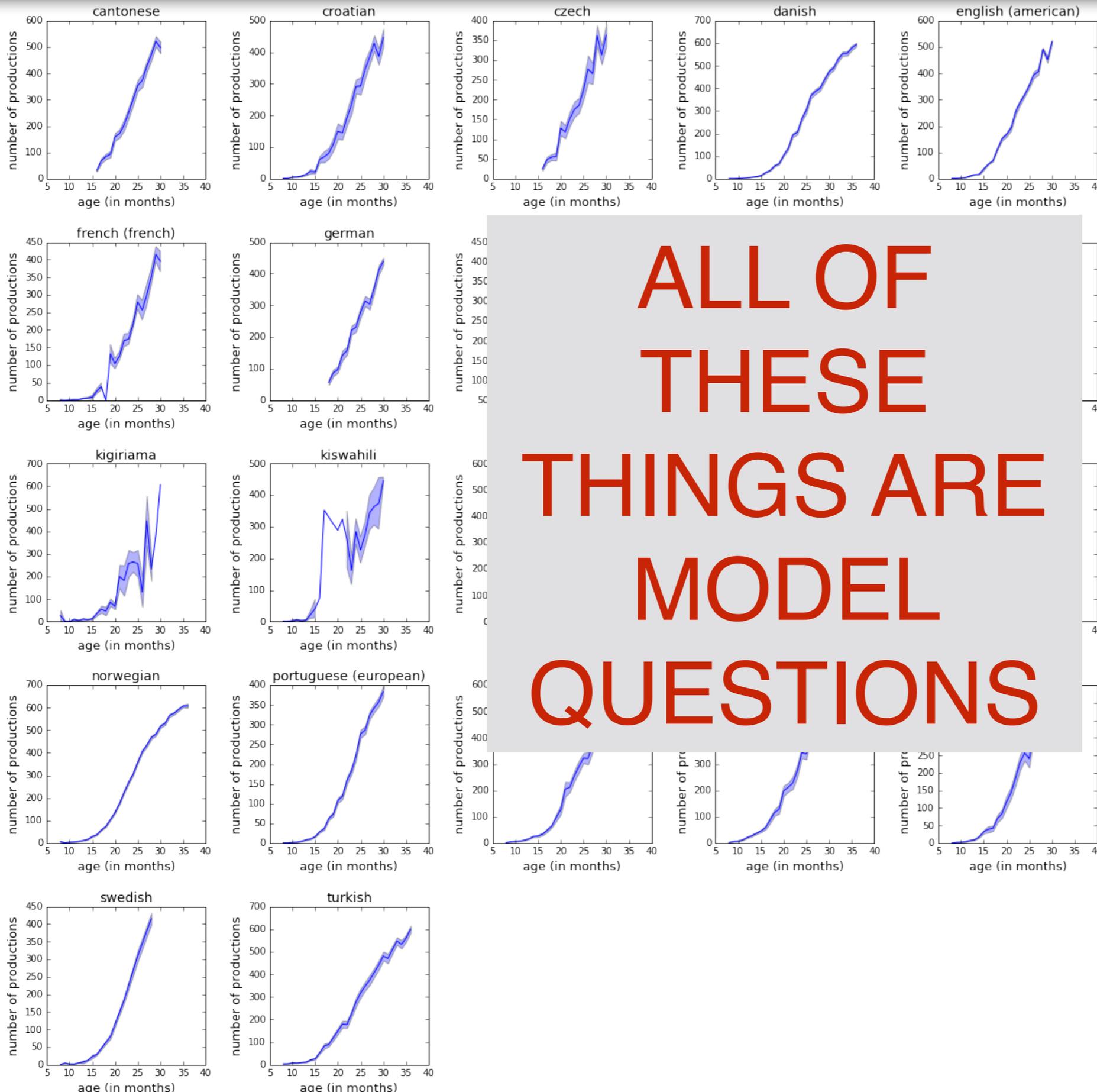
Is the functional form the same across languages?

Assuming they're all sigmoids with different parameters, what is the distribution of parameters?

How should they be interpreted?

What causes them to vary?

Example 3: The WordBank Dataset



Is the functional form the same across languages?

Assuming they're all sigmoids with different parameters, what is the distribution of parameters?

How should they be interpreted?

What causes them to vary?

Main Goals of Class

Learn how to

- 1) Build (high-dimensional) models of/for neural & behavioral data

$$\text{output} = F_{[\text{params}]}(\text{input})$$

*formulate **F** mathematically
on on a computer*



Main Goals of Class

Learn how to

1) Build (high-dimensional) models of/for neural & behavioral data

$$\text{output} = F_{[\text{params}]}(\text{input})$$

*formulate **F** mathematically
on on a computer*



2) Train such models

$$\text{data_slice_2} = F_{[\text{params}]}(\text{data_slice_1})$$

*determine **params**
from pair of linked
data tensors*



Main Goals of Class

Learn how to

1) Build (high-dimensional) models of/for neural & behavioral data

$$\text{output} = F_{[\text{params}]}(\text{input})$$

*formulate **F** mathematically
on on a computer*

2) Train such models

$$\text{data_slice_2} = F_{[\text{params}]}(\text{data_slice_1})$$

*determine **params**
from pair of linked
data tensors*

3) Evaluate and compare such models

$$\text{new_data_slice_2} = F_{[\text{params}]}(\text{new_data_slice_1})$$

Main Goals of Class

Supervised training:

$$\text{data_slice_2} = F_{[\text{params}]}(\text{data_slice_1})$$

determine **params**
from pair of linked
data tensors



parameter-finding procedure involves some examples of data_slice_2,
evaluation has to be done on held-out data

Main Goals of Class

Unsupervised training:

exploratory idea = $F_{[\text{params}]}(\text{data_slice_I})$ *determine **params** from single tensor*



parameter-finding procedure doesn't involve pre-determined output

Task performance data

Survey responses

Health outcomes

Table 1
Cognitive tasks.

Task	Individual Difference Variables	References
Adaptive N-Back	DDM Parameters Average load	(Harvey et al., 2005; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008)
Angling Risk Task	Two Conditions (Keep, Release): Adjusted Clicks Loss Percent Score	(Pleskac, 2008)
Attention Network Task	DDM Parameters Alerting Effect Orient Effect Conflict Effect	(Fan, McCandliss, Fossella, Flombaum, & Posner, 2005)
Bickel Titrator	Discount Rate for three payout magnitudes	(Koffarnus & Bickel, 2014)
Choice Reaction Time Cognitive Reflection Task	DDM Parameters Correct Proportion Intuitive Proportion	(Primi, Morsanyi, Chiesi, Donati, & Hamilton, 2016; Toplak, West, & Stanovich, 2014)
Columbia Card Task Cold/Hot	Average # of cards chosen Gain Sensitivity Loss Sensitivity # Loss Cards Sensitivity Level of Information Use	(Figner, Mackinlay, Wilkening, & Weber, 2009)
Dietary Decision Task	Health Sensitivity Taste Sensitivity	(Hare, Camerer, & Rangel, 2009)
Digit Span	Forward Span Reverse Span	(Woods et al., 2011)
Directed Forgetting	DDM Parameters Proactive Interference	(Nee, Jonides, & Berman, 2007)
Discount Titrator	Percent Patient	(Figner et al., 2010)
Dot Pattern Expectancy	DDM Parameters AY-BY BX-BY D-prime Bias	(Otto, Skatova, Madlon-Kay, & Daw, 2013)

Table 2
Self-report surveys.

Self-Report Surveys	Individual Difference Variables	References
BIS-11	Attentional Motor Non-Planning	(Patton, Stanford, & Barratt, 1995)
BIS-BAS	BAS Drive BAS Fun-Seeking BAS Reward-Responsiveness BIS	(Carver & White, 1994)
Brief Self-Control Scale	Self-Control	(Roth, Isquith, & Gioia, 2005) (Dickman, 1990)
Dickman's Impulsivity Inventory DOSPERT (EB/RP/RT)	Dysfunctional Functional Ethical Financial Health/Safety Recreational Social	(Blais & Weber, 2006)
Three-Factor Eating Questionnaire (R18)	Cognitive Restraint Emotional Eating Uncontrolled Eating	(de Lauzon et al., 2004)
Emotion Regulation Questionnaire	Reappraisal Suppression	(Gross & John, 2003)
Five Facet Mindfulness Questionnaire	Acts with Awareness Describe Non-Judgment Non-Reactive Observe	(Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006)
Future Time Perspective	Future-Time Perspective	(Carstensen & Lang, 1996)
Grit Scale	Grit	(Duckworth and Quinn, 2009)

Demographics

Sex	Age	Race	OtherRace	HispanicLatino	HighestEducation	HeightInches	WeightPounds	RelationshipStatus
DivorceCount	LongestRelationship	RelationshipNumber	ChildrenNumber	HouseholdIncome	RetirementAccount	RetirementPercentStocks	RentOwn	
MortgageDebt	CarDebt	EducationDebt	CreditCardDebt	OtherDebtSources		OtherDebtAmount		
CoffeeCupsPerDay	TeaCupsPerDay	CaffienatedSodaCansPerDay	CaffieneOtherSourcesDayMG	GamblingProblem	TrafficTicketsLastYearCount			
TrafficAccidentsLifeCount	ArrestedChargedLifeCount	MotivationForParticipation						

NeurologicalDiagnoses
NeurologicalDiagnosesDescribe

DoctorVisitsLastMonth	Nervous
DaysPhysicalHealthFeelings	Hopeless
PsychDiagnoses	RestlessFidgety
PsychDiagnoses	Depressed
PsychDiagnosesDescribe	EverythingIsEffort
PsychDiagnosesDescribe	Worthless
PsychDiagnosesDescribe	Last30DaysUsual
PsychDiagnosesDescribe	DaysLostLastMonth
PsychDiagnosesDescribe	DaysHalfLastMonth

Task performance data

Table 1
Cognitive tasks.

Task	Individual Difference Variables	References
Adaptive N-Back	DDM Parameters Average load	(Harvey et al., 2005; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008)
Angling Risk Task	Two Conditions (Keep, Release): Adjusted Clicks Loss Percent Score	(Pleskac, 2008)
Attention Network Task	DDM Parameters Alerting Effect Orient Effect Conflict Effect	(Fan, McCandliss, Fossella, Flombaum, & Posner, 2005)
Bickel Titrator	Discount Rate for three payout magnitudes	(Koffarnus & Bickel, 2014)
Choice Reaction Time Cognitive Reflection Task	DDM Parameters Correct Proportion Intuitive Proportion	(Primi, Morsanyi, Chiesi, Donati, & Hamilton, 2016; Toplak, West, & Stanovich, 2014)
Duration of Card Chosen Cold/Hot	Average of Card chosen Gain Sensitivity Loss Sensitivity # Loss Cards Sensitivity Level of Information Use	(Figner, Mackinlay, Wilkening, & Weber, 2009)
Dietary Decision Task	Health Sensitivity Taste Sensitivity	(Hare, Camerer, & Rangel, 2009)
Digit Span	Forward Span Reverse Span	(Cotton et al., 2011)
Directed Forgetting	DDM Parameters Proactive Interference	(Lee, Jonides, & Berman, 2007)
Discount Titrator	Percent Patient	(Figner et al., 2010)
Dot Pattern Expectancy	DDM Parameters AY-BY BX-BY D-prime Bias	(Otto, Skatova, Madlon-Kay, & Law, 2013)

Survey responses

Table 2
Self-report surveys.

Self-Report Surveys	Individual Difference Variables	References
BIS-11	Attentional Motor Non-Planning	(Patton, Stanford, & Barratt, 1995)
BIS-BAS	BAS Drive BAS Fun-Seeking BAS Reward-Responsiveness BIS	(Carver & White, 1994)
Brief Self-Control Scale	Self-Control	(Roth, Isquith, & Gioia, 2005)
Dickman's Impulsivity Inventory	Dysfunctional Functional	(Dickman, 1990)
DOSPERS (EB/RP/RT)	Ethical Financial Health/Safety Recreational Social	(Blais & Weber, 2006)
Three-Factor Eating Questionnaire (R18)	Cognitive Restraint Emotional Eating Uncontrolled Eating	(de Lauzon et al., 2004)
Emotion Regulation Questionnaire	Reappraisal Suppression	(Gross & John, 2003)
Five Facet Mindfulness Questionnaire	Acts with Awareness Describe Non-Judgment Non-Reactive Observe	(Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006)
Future Time Perspective	Future-Time Perspective	(Carstensen & Lang, 1996)
Grit Scale	Grit	(Duckworth and Quinn, 2009)

Health outcomes

NeurologicalDiagnoses
NeurologicalDiagnosesDescribe

DoctorVisitsLastMonth	Nervous
DaysPhysicalHealthFeelings	Hopeless
PsychDiagnoses	RestlessFidgety
PsychDiagnoses	Depressed
PsychDiagnosesDescribe	EverythingIsEffort
PsychDiagnosesDescribe	Worthless
PsychDiagnosesDescribe	Last30DaysUsual
PsychDiagnosesDescribe	DaysLostLastMonth
PsychDiagnosesDescribe	DaysHalfLastMonth

Demographics

Sex	Age	Race	OtherRace	HispanicLatino	HighestEducation	HeightInches	WeightPounds	RelationshipStatus
DivorceCount	LongestRelationship	RelationshipNumber	ChildrenNumber	HouseholdIncome	RetirementAccount	RetirementPercentStocks	RentOwn	
MortgageDebt	CarDebt	EducationDebt	CreditCardDebt	OtherDebtSources	OtherDebtAmount			
CoffeeCupsPerDay	TeaCupsPerDay	CaffienatedSodaCansPerDay	CaffieneOtherSourcesDayMG	GamblingProblem	TrafficTicketsLastYearCount			
TrafficAccidentsLifeCount	ArrestedChargedLifeCount	MotivationForParticipation						

Task performance data

Survey responses

Health outcomes

Table 1
Cognitive tasks.

Task	Individual Difference Variables	References
Adaptive N-Back	DDM Parameters Average load	(Harvey et al., 2005; Jaeggi, Buschkuhl, Jonides, & Perrig, 2008)
Angling Risk Task	Two Conditions (Keep, Release): Adjusted Clicks Loss Percent Score	(Pleskac, 2008)
Attention Network Task	DDM Parameters Alerting Effect Orient Effect Conflict Effect	(Fan, McCandliss, Fossella, Flombaum, & Posner, 2005)
Bickel Titrator	Discount Rate for three payout magnitudes	(Koffarnus & Bickel, 2014)
Choice Reaction Time Cognitive Reflection Task	DDM Parameters Correct Proportion Intuitive Proportion	(Primi, Morsanyi, Chiesi, Donati, & Hamilton, 2016; Toplak, West, & Stanovich, 2014)
Duration of Task Cold/Hot	Average of 4rd chosen Gain Sensitivity Loss Sensitivity # Loss Cards Sensitivity Level of Information Use	(Figner, Mackinlay, Wilkening, & Weber, 2009)
Dietary Decision Task	Health Sensitivity Taste Sensitivity	(Hare, Camerer, & Rangel, 2009)
Digit Span	Forward Span Reverse Span	(Cotton et al., 2011)
Directed Forgetting	DDM Parameters Proactive Interference	(Lee, Jonides, & Berman, 2007)
Discount Titrator	Percent Patient	(Figner et al., 2010)
Dot Pattern Expectancy	DDM Parameters AY-BY BX-BY D-prime Bias	(Otto, Skatova, Madlon-Kay, & Law, 2013)

Table 2
Self-report surveys.

Self-Report Surveys	Individual Difference Variables	References
BIS-11	Attentional Motor Non-Planning	(Patton, Stanford, & Barratt, 1995)
BIS-BAS	BAS Drive BAS Fun-Seeking BAS Reward-Responsiveness	(Carver & White, 1994)
Brief Self-Control Scale	BIS Self-Control	(Roth, Isquith, & Gioia, 2005)
Dickman's Impulsivity Inventory	Dysfunctional Functional	(Dickman, 1990)
DOSPERT (EB/RP/RT)	Ethical Financial Health/Safety Recreational Social	(Blais & Weber, 2006)
Three-Factor Eating Questionnaire (R18)	Cognitive Restraint Emotional Eating Uncontrolled Eating	(de Lauzon et al., 2004)
Emotion Regulation Questionnaire	Reappraisal Suppression	(Gross & John, 2003)
Five Facet Mindfulness Questionnaire	Acts with Awareness Describe Non-Judgment Non-Reactive Observe	(Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006)
Future Time Perspective	Future-Time Perspective	(Carstensen & Lang, 1996)
Grit Scale	Grit	(Duckworth and Quinn, 2009)

Discovery of natural factors & survey design

Demographics

Sex	Age	Race	OtherRace	HispanicLatino	HighestEducation	HeightInches	WeightPounds	RelationshipStatus
DivorceCount	LongestRelationship	RelationshipNumber	ChildrenNumber	HouseholdIncome	RetirementAccount	RetirementPercentStocks	RentOwn	
MortgageDebt	CarDebt	EducationDebt	CreditCardDebt	OtherDebtSources	OtherDebtAmount			
CoffeeCupsPerDay	TeaCupsPerDay	CaffienatedSodaCansPerDay	CaffieneOtherSourcesDayMG	GamblingProblem	TrafficTicketsLastYearCount			
TrafficAccidentsLifeCount	ArrestedChargedLifeCount	MotivationForParticipation						

NeurologicalDiagnoses	DoctorVisitsLastMonth	Nervous
	DaysPhysicalHealthFeelings	Hopeless
PsychDiagnoses	DaysPhysicalHealthFeelings	RestlessFidgety
	PsychDiagnoses	Depressed
NeurologicalDiagnosesDescribe	PsychDiagnoses	EverythingIsEffort
	NeurologicalDiagnosesDescribe	Worthless
PsychDiagnosesOther	PsychDiagnosesDescribe	Last30DaysUsual
	PsychDiagnosesOther	DaysLostLastMonth
		DaysHalfLastMonth

The Four Ingredients of Any Model

ingredient 0

The data:

Observations used to determine
the model & to evaluate it

The Four Ingredients of Any Model

ingredient 0

The data:

Observations used to determine the model & to evaluate it

ingredient 1

The model type:

parameterized
function class

The Four Ingredients of Any Model

ingredient 0

The data:

Observations used to determine the model & to evaluate it

ingredient 1

The model type:

parameterized function class

ingredient 2

The loss function:

what do you want the model to do?

The Four Ingredients of Any Model

ingredient 0

The data:
Observations used to determine
the model & to evaluate it

ingredient 1

The model type:
parameterized
function class

ingredient 2

The loss function:
what do you want
the model to do?

ingredient 3

The optimization procedure:
how are you going to estimate the
parameters from your data?

ingredient 1

The model type:
parameterized
function class

example: **linear classifiers**

linear + threshold

ingredient 2

The loss function:
what do you want
the model to do?

svm: hinge loss
logistic: maximum likelihood

ingredient 3

The optimization procedure:
how are you going to estimate the
parameters from your data?

Gradient descent

ingredient 1

The model type:
parameterized
function class

example: **k-means clustering**

discrete partition

ingredient 2

The loss function:
what do you want
the model to do?

in-vs-out-group variance

ingredient 3

The optimization procedure:
how are you going to estimate the
parameters from your data?

lloyd's algorithm

The Cross-Validation Framework For the Four Ingredients of Any Model

Cross-validation: estimate then test on held-out data

ingredient 1

The model type:
parameterized
function class

ingredient 2

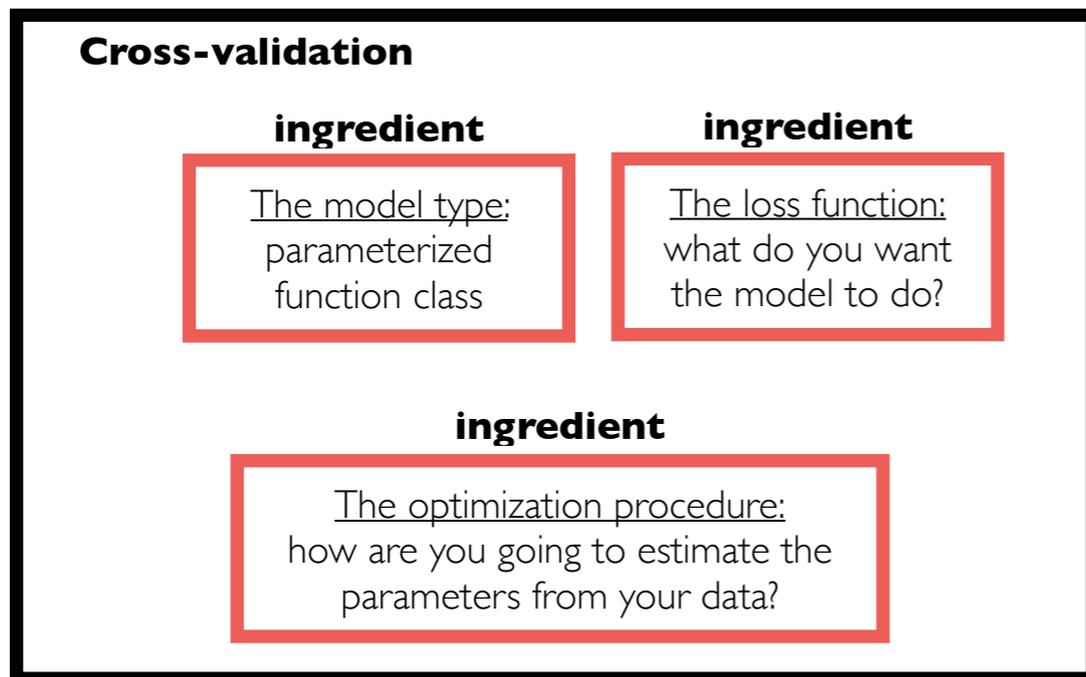
The loss function:
what do you want
the model to do?

ingredient 3

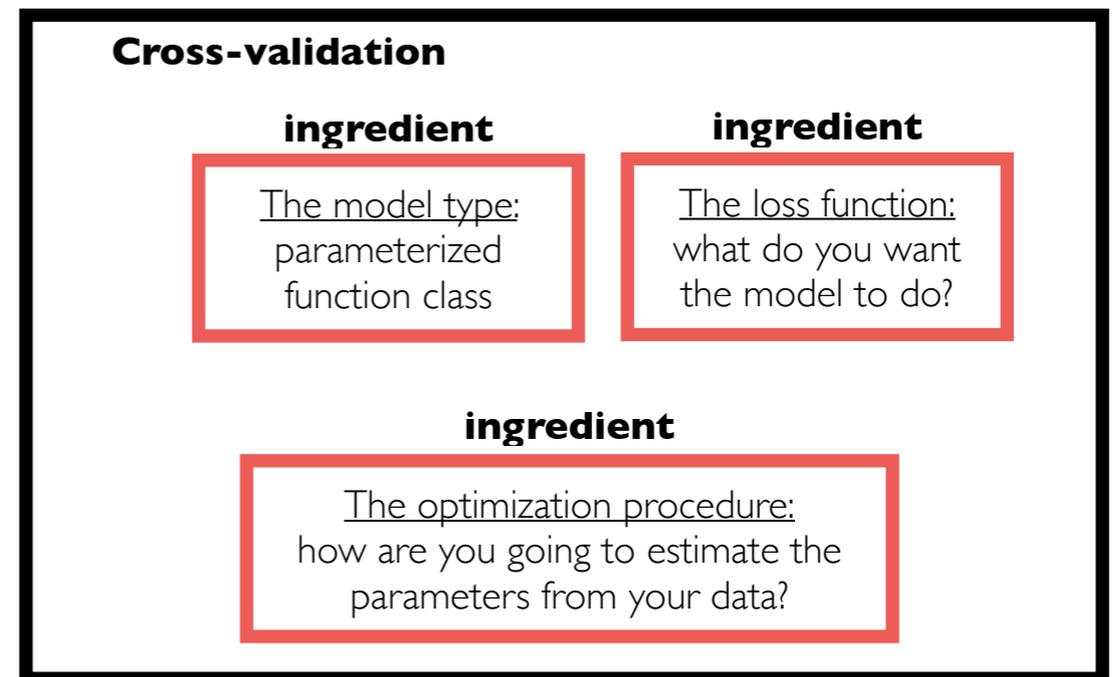
The optimization procedure:
how are you going to estimate the
parameters from your data?

Then you compare model correctness

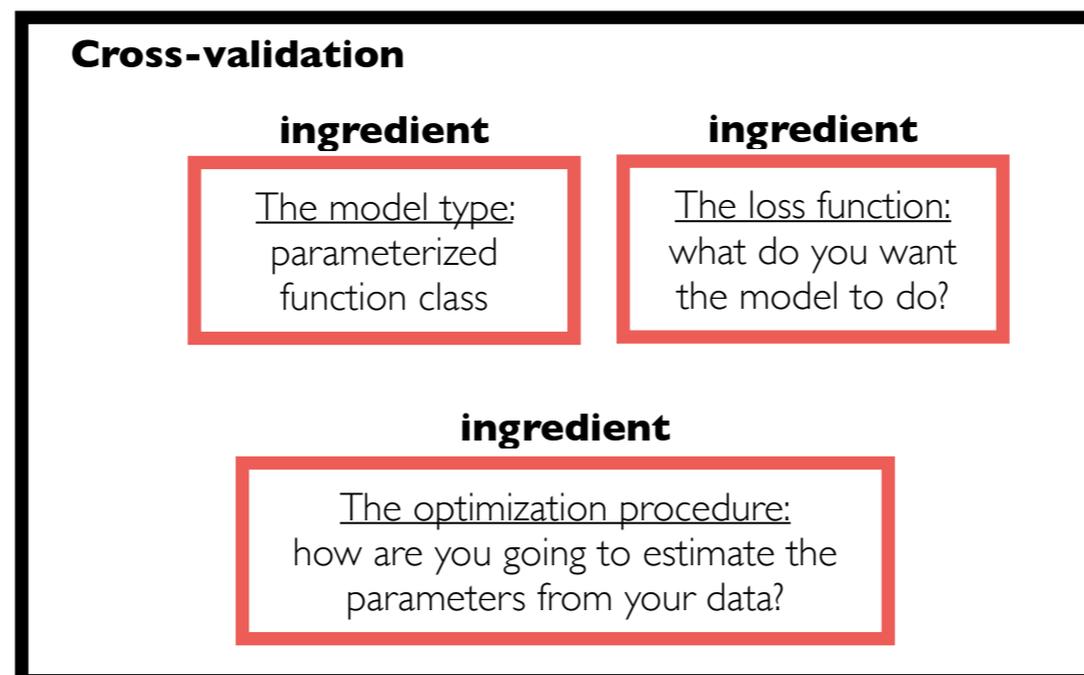
model 1



model 2



model 3

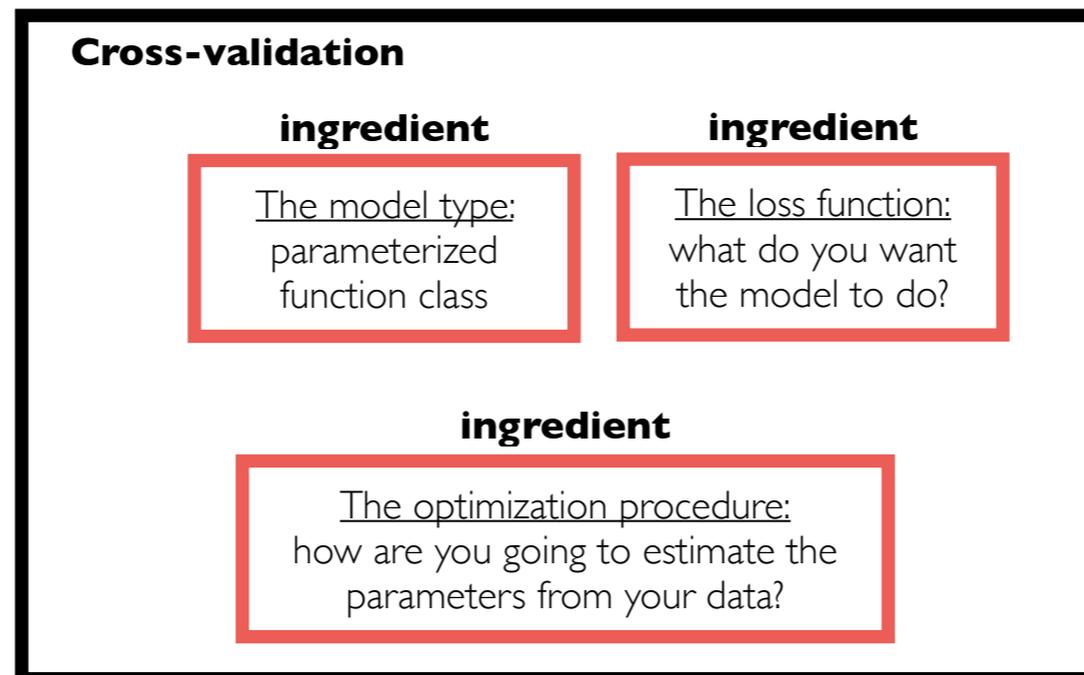
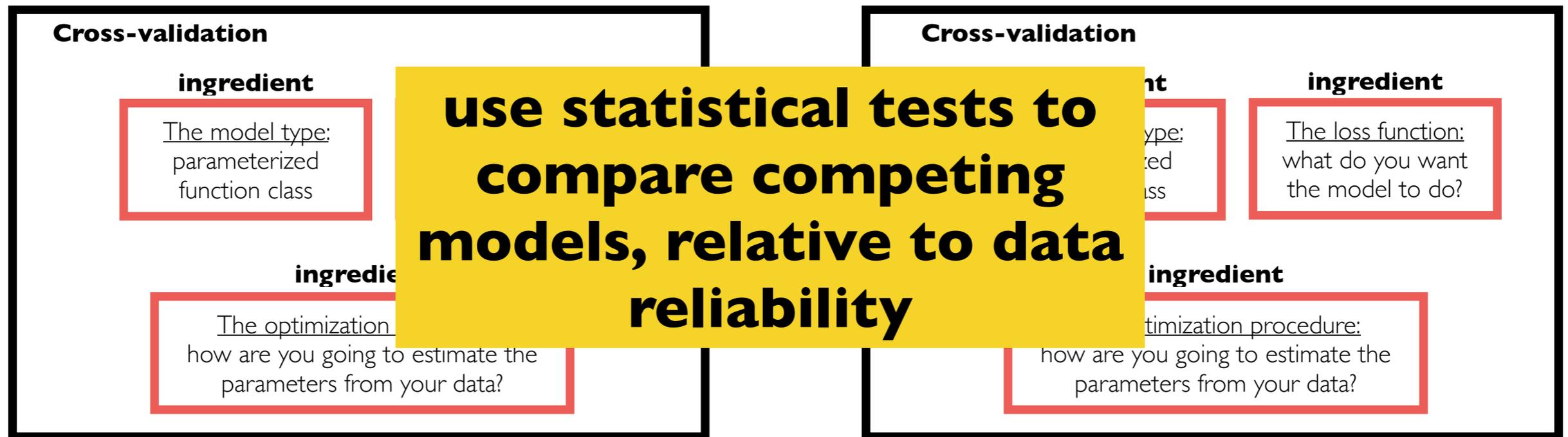


The Unified Four-Ingredient Cross-Validated Model-Comparison Framework

Then you compare model correctness

model 1

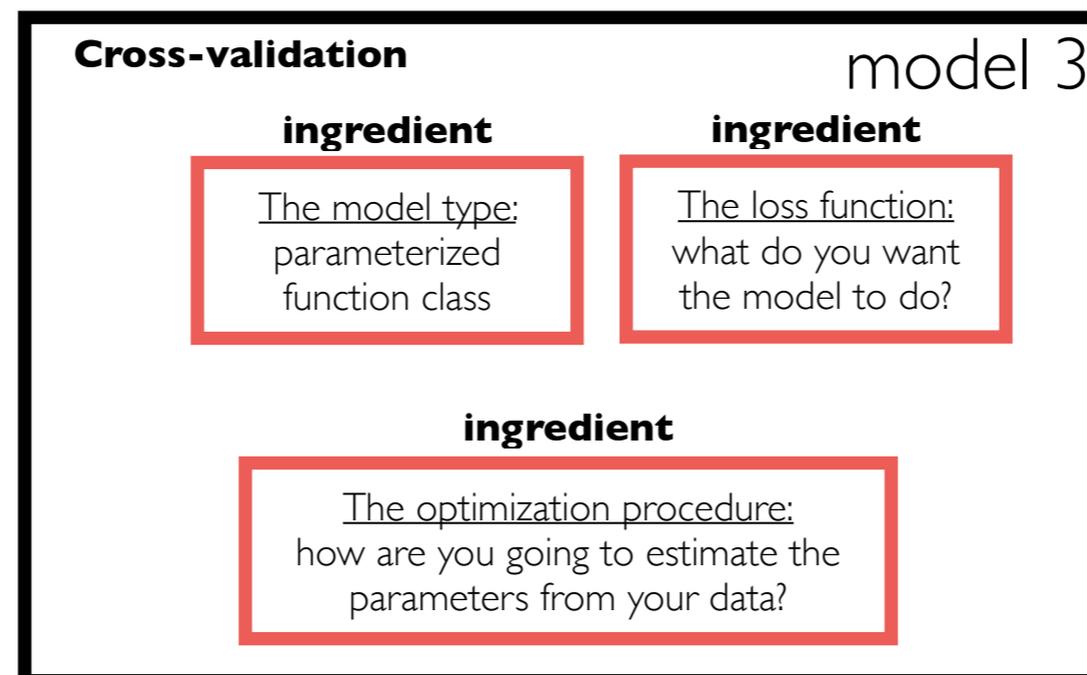
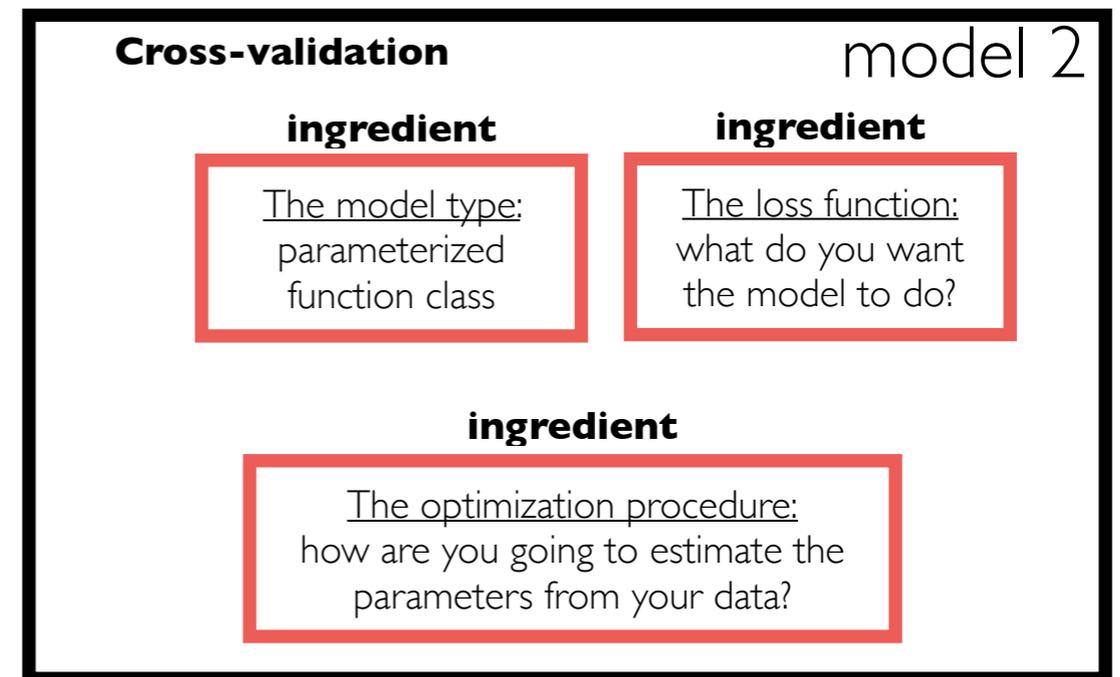
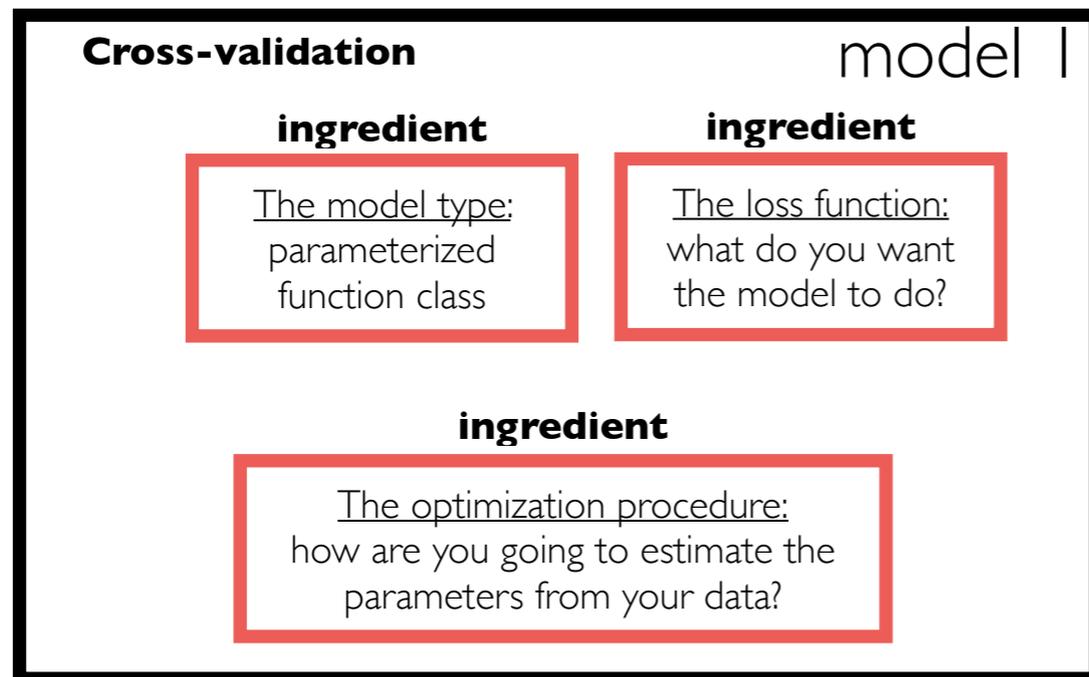
model 2



model 3

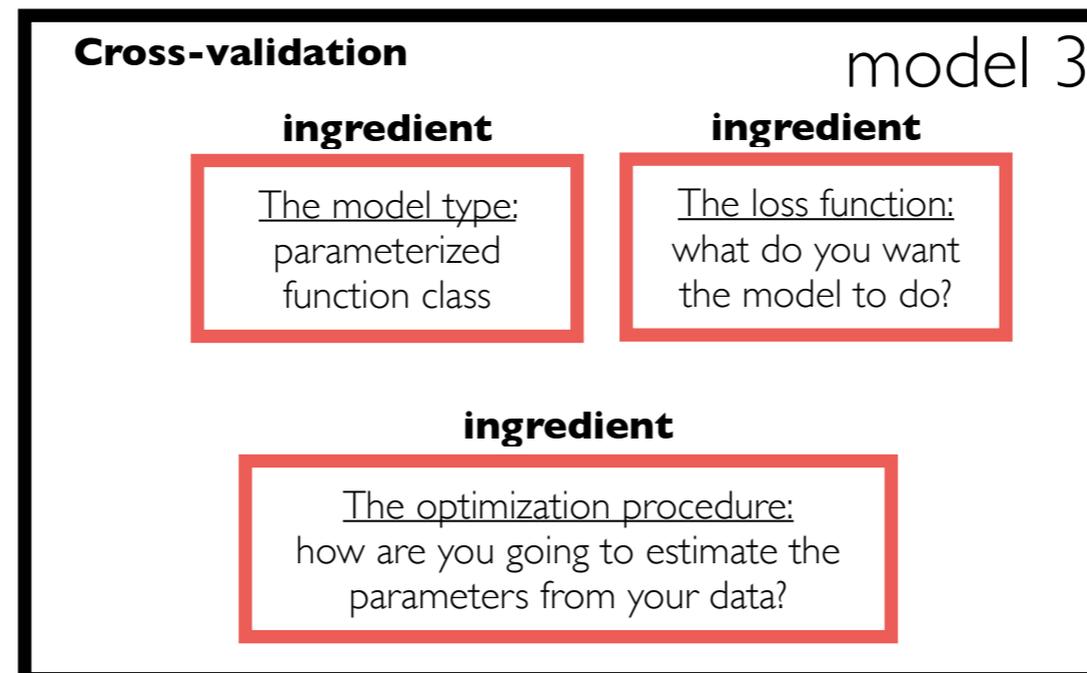
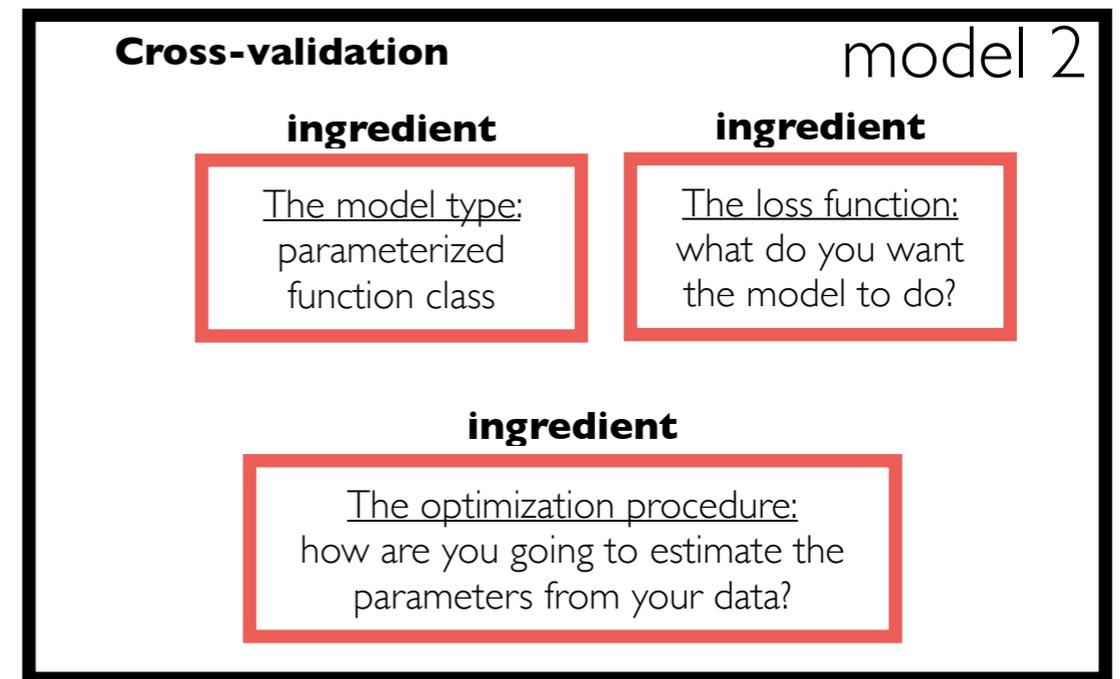
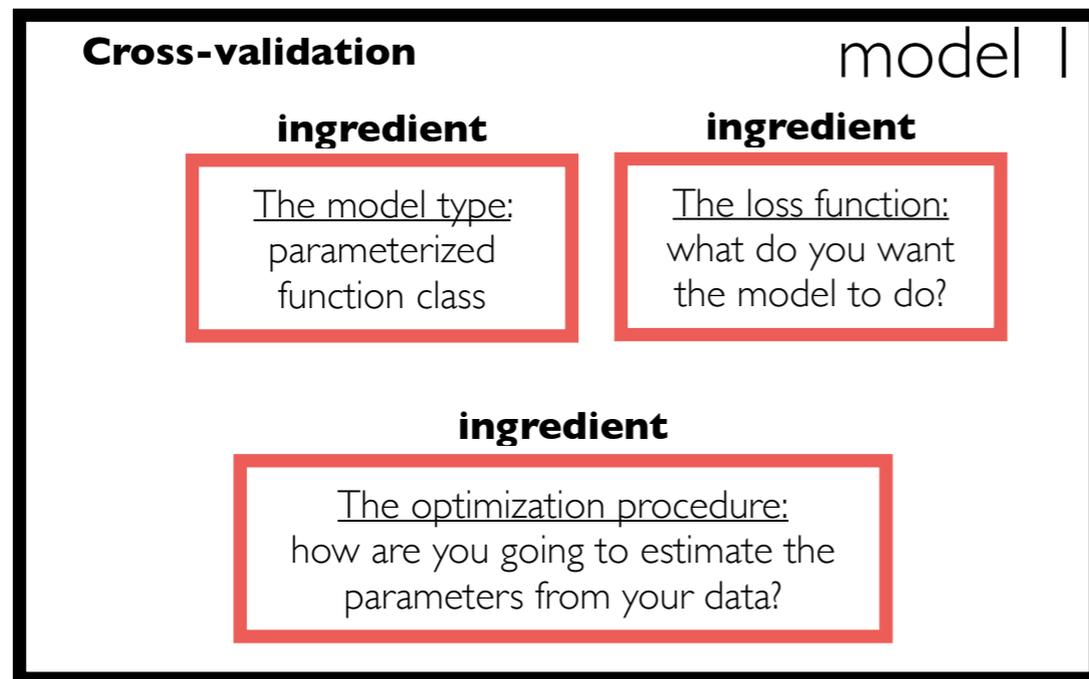
Then you compare model correctness

Scientific enterprise



key experimental design principle: your data should support differentiating between models

Scientific enterprise



Class Schedule

Date	Session	ipy nb	Assignments
03/29	<u>Foundations</u> What is Data? What are Models?	<u>Lecture 1</u>	<u>HW1 Released</u>
03/31	<u>Foundations</u> Modeling: a framework and strategy		
04/05	<u>Foundations</u> Reliability: the data as a model of itself		
04/07	<u>Foundations</u> Regression: OLS		
04/12	<u>Foundations</u> Correlation Analysis		
04/14	<u>Supervised Models</u> Minimum distance classifiers		HW1 Due; <u>HW2 released</u>
04/19	<u>Supervised Models</u> SVMs and logistic regression		
04/21	<u>Supervised Models</u> Regularization		
04/26	<u>Supervised Models</u> Mixed effects and hierarchical models		Project description due
04/28	<u>Unsupervised Models</u> Clustering		
05/03	<u>Unsupervised Models</u> Linear Algebra Review		
05/05	<u>Unsupervised Models</u> Dimension Reduction		HW2 Due
05/10	<u>Timeseries Models</u> Guest lecture from Nilam Ram		
05/12	<u>Causal Models</u> Structural Equation Modeling		
05/17	<u>Causal Models</u> Graphical Models		
05/19	<u>Causal Models</u> Network Analysis		
05/24	<u>Foundations</u> Optimization		
05/26	<u>Foundations</u> Making Your Own Custom Models		
05/31	Memorial Day Holiday - no class		
06/02	Project Presentations		Project writeup due

Class Schedule

Date	Session	ipy nb	Assignments
03/29	<u>Foundations</u> What is Data? What are Models?	<u>Lecture 1</u>	<u>HW1 Released</u>
03/31	<u>Foundations</u> Modeling: a framework and strategy		
04/05	<u>Foundations</u> Reliability: the data as a model of itself		
04/07	<u>Foundations</u> Regression: OLS		
04/12	<u>Foundations</u> Correlation Analysis		
04/14	<u>Supervised Models</u> Minimum distance classifiers		HW1 Due; <u>HW2 released</u>
04/19	<u>Supervised Models</u> SVMs and logistic regression		
04/21	<u>Supervised Models</u> Regularization		
04/26	<u>Supervised Models</u> Mixed effects and hierarchical models		Project description due
04/28	<u>Unsupervised Models</u> Clustering		
05/03	<u>Unsupervised Models</u> Linear Algebra Review		
05/05	<u>Unsupervised Models</u> Dimension Reduction		HW2 Due
05/10	<u>Timeseries Models</u> Guest lecture from Nilam Ram		
05/12	<u>Causal Models</u> Structural Equation Modeling		
05/17	<u>Causal Models</u> Graphical Models		
05/19	<u>Causal Models</u> Network Analysis		
05/24	<u>Foundations</u> Optimization		
05/26	<u>Foundations</u> Making Your Own Custom Models		
05/31	Memorial Day Holiday - no class		
06/02	Project Presentations		Project writeup due

Class Schedule

Date	Session	ipy nb	Assignments
03/29	<u>Foundations</u> What is Data? What are Models?	<u>Lecture 1</u>	<u>HW1 Released</u>
03/31	<u>Foundations</u> Modeling: a framework and strategy		
04/05	<u>Foundations</u> Reliability: the data as a model of itself		
04/07	<u>Foundations</u> Regression: OLS		
04/12	<u>Foundations</u> Correlation Analysis		
04/14	<u>Supervised Models</u> Minimum distance classifiers		HW1 Due; <u>HW2 released</u>
04/19	<u>Supervised Models</u> SVMs and logistic regression		
04/21	<u>Supervised Models</u> Regularization		
04/26	<u>Supervised Models</u> Mixed effects and hierarchical models		Project description due
04/28	<u>Unsupervised Models</u> Clustering		
05/03	<u>Unsupervised Models</u> Linear Algebra Review		
05/05	<u>Unsupervised Models</u> Dimension Reduction		HW2 Due
05/10	<u>Timeseries Models</u> Guest lecture from Nilam Ram		
05/12	<u>Causal Models</u> Structural Equation Modeling		
05/17	<u>Causal Models</u> Graphical Models		
05/19	<u>Causal Models</u> Network Analysis		
05/24	<u>Foundations</u> Optimization		
05/26	<u>Foundations</u> Making Your Own Custom Models		
05/31	Memorial Day Holiday - no class		
06/02	Project Presentations		Project writeup due

Class Schedule

Date	Session	ipy nb	Assignments
03/29	<u>Foundations</u> What is Data? What are Models?	<u>Lecture 1</u>	<u>HW1 Released</u>
03/31	<u>Foundations</u> Modeling: a framework and strategy		
04/05	<u>Foundations</u> Reliability: the data as a model of itself		
04/07	<u>Foundations</u> Regression: OLS		
04/12	<u>Foundations</u> Correlation Analysis		
04/14	<u>Supervised Models</u> Minimum distance classifiers		HW1 Due; <u>HW2 released</u>
04/19	<u>Supervised Models</u> SVMs and logistic regression		
04/21	<u>Supervised Models</u> Regularization		
04/26	<u>Supervised Models</u> Mixed effects and hierarchical models		Project description due
04/28	<u>Unsupervised Models</u> Clustering		
05/03	<u>Unsupervised Models</u> Linear Algebra Review		
05/05	<u>Unsupervised Models</u> Dimension Reduction		HW2 Due
05/10	<u>Timeseries Models</u> Guest lecture from Nilam Ram		
05/12	<u>Causal Models</u> Structural Equation Modeling		
05/17	<u>Causal Models</u> Graphical Models		
05/19	<u>Causal Models</u> Network Analysis		
05/24	<u>Foundations</u> Optimization		
05/26	<u>Foundations</u> Making Your Own Custom Models		
05/31	Memorial Day Holiday - no class		
06/02	Project Presentations		Project writeup due

Class Schedule

Date	Session	ipy nb	Assignments
03/29	<u>Foundations</u> What is Data? What are Models?	<u>Lecture 1</u>	<u>HW1 Released</u>
03/31	<u>Foundations</u> Modeling: a framework and strategy		
04/05	<u>Foundations</u> Reliability: the data as a model of itself		
04/07	<u>Foundations</u> Regression: OLS		
04/12	<u>Foundations</u> Correlation Analysis		
04/14	<u>Supervised Models</u> Minimum distance classifiers		HW1 Due; <u>HW2 released</u>
04/19	<u>Supervised Models</u> SVMs and logistic regression		
04/21	<u>Supervised Models</u> Regularization		
04/26	<u>Supervised Models</u> Mixed effects and hierarchical models		Project description due
04/28	<u>Unsupervised Models</u> Clustering		
05/03	<u>Unsupervised Models</u> Linear Algebra Review		
05/05	<u>Unsupervised Models</u> Dimension Reduction		HW2 Due
05/10	<u>Timeseries Models</u> Guest lecture from Nilam Ram		
05/12	<u>Causal Models</u> Structural Equation Modeling		
05/17	<u>Causal Models</u> Graphical Models		
05/19	<u>Causal Models</u> Network Analysis		
05/24	<u>Foundations</u> Optimization		
05/26	<u>Foundations</u> Making Your Own Custom Models		
05/31	Memorial Day Holiday - no class		
06/02	Project Presentations		Project writeup due

Class Schedule

Date	Session	ipy nb	Assignments
03/29	<u>Foundations</u> What is Data? What are Models?	<u>Lecture 1</u>	<u>HW1 Released</u>
03/31	<u>Foundations</u> Modeling: a framework and strategy		
04/05	<u>Foundations</u> Reliability: the data as a model of itself		
04/07	<u>Foundations</u> Regression: OLS		
04/12	<u>Foundations</u> Correlation Analysis		
04/14	<u>Supervised Models</u> Minimum distance classifiers		HW1 Due; <u>HW2 released</u>
04/19	<u>Supervised Models</u> SVMs and logistic regression		
04/21	<u>Supervised Models</u> Regularization		
04/26	<u>Supervised Models</u> Mixed effects and hierarchical models		Project description due
04/28	<u>Unsupervised Models</u> Clustering		
05/03	<u>Unsupervised Models</u> Linear Algebra Review		
05/05	<u>Unsupervised Models</u> Dimension Reduction		HW2 Due
05/10	<u>Timeseries Models</u> Guest lecture from Nilam Ram		
05/12	<u>Causal Models</u> Structural Equation Modeling		
05/17	<u>Causal Models</u> Graphical Models		
05/19	<u>Causal Models</u> Network Analysis		
05/24	<u>Foundations</u> Optimization		
05/26	<u>Foundations</u> Making Your Own Custom Models		
05/31	Memorial Day Holiday - no class		
06/02	Project Presentations		Project writeup due

Class Schedule

Date	Session	ipy nb	Assignments
03/29	<u>Foundations</u> What is Data? What are Models?	Lecture 1	HW1 Released
03/31	<u>Foundations</u> Modeling: a framework and strategy		
04/05	<u>Foundations</u> Reliability: the data as a model of itself		
04/07	<u>Foundations</u> Regression: OLS		
04/12	<u>Foundations</u> Correlation Analysis		
04/14	<u>Supervised Models</u> Minimum distance classifiers		HW1 Due; HW2 released
04/19	<u>Supervised Models</u> SVMs and logistic regression		
04/21	<u>Supervised Models</u> Regularization		
04/26	<u>Supervised Models</u> Mixed effects and hierarchical models		Project description due
04/28	<u>Unsupervised Models</u> Clustering		
05/03	<u>Unsupervised Models</u> Linear Algebra Review		
05/05	<u>Unsupervised Models</u> Dimension Reduction		HW2 Due
05/10	<u>Timeseries Models</u> Guest lecture from Nilam Ram		
05/12	<u>Causal Models</u> Structural Equation Modeling		
05/17	<u>Causal Models</u> Graphical Models		
05/19	<u>Causal Models</u> Network Analysis		
05/24	<u>Foundations</u> Optimization		
05/26	<u>Foundations</u> Making Your Own Custom Models		
05/31	Memorial Day Holiday - no class		
06/02	Project Presentations		Project writeup due

Main Goals of Class

**Logistics & announcements
& discussion:**

stanfordpsych253.slack.com

**Schedule & syllabus &
Slides and notebooks:**

psych253.stanford.edu