What's wrong with models of ventral temporal cortex?

UCLA, July 2016

Jim Haxby
Center for Cognitive Neuroscience, Dartmouth College
Center for Mind/Brain Sciences (CIMeC), University of Trento

CCN Center for Cognitive Neuroscience at Dartmouth

CIMeC

Neural representation of faces, bodies, and objects in ventral temporal cortex

UCLA, July 2016

Jim Haxby
Center for Cognitive Neuroscience, Dartmouth College
Center for Mind/Brain Sciences (CIMeC), University of Trento

CCN Center for Cognitive Neuroscience at Dartmouth

CIMeC

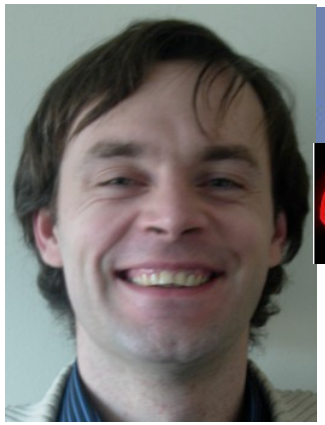
Students, postdocs, collaborators

Analysis of similarity structure,
representation of biological classes
Andy Connolly
Research assistant professor
(Geisel School of Medicine)

Hyperalignment
Swaroop Guntupalli
Post-doctoral fellow

Attention
Sam Nastase
Graduate student

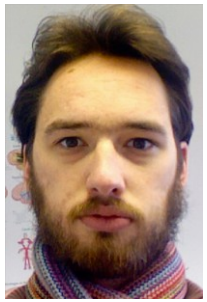
Person perception
Dylan Wagner
Assistant professor
(Ohio State)



PyMVPA

NeuroDebian

Yaroslav Halchenko
Research scientist



Action representation, computational methods
Nick Oosterhof
Post-doctoral fellow (CIMeC)

Neural representation of faces, bodies, and objects in ventral temporal cortex

Optimized for person perception

- Face processing is fast and requires minimal attention and awareness
- Mediated by a distributed neural system in humans and monkeys

A fly in the ointment

- Stimulus sampling bias

  ? Representation of agentic action plays a dominant role
- Inadequate computational methods for data analysis and modeling

  ? Modeling functional architecture as a high-dimensional space captures complexity better than category-specific regions

Neural representation of faces, bodies, and objects in ventral temporal cortex

Optimized for person perception

- Face processing is fast and requires minimal attention and awareness
- Mediated by a distributed neural system in humans and monkeys

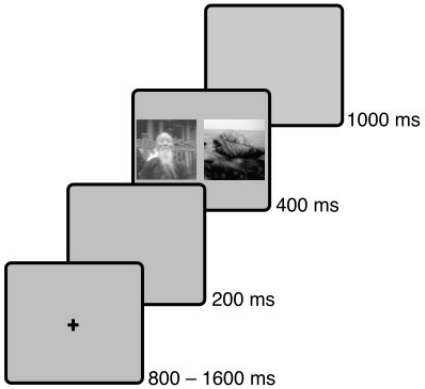
A fly in the ointment

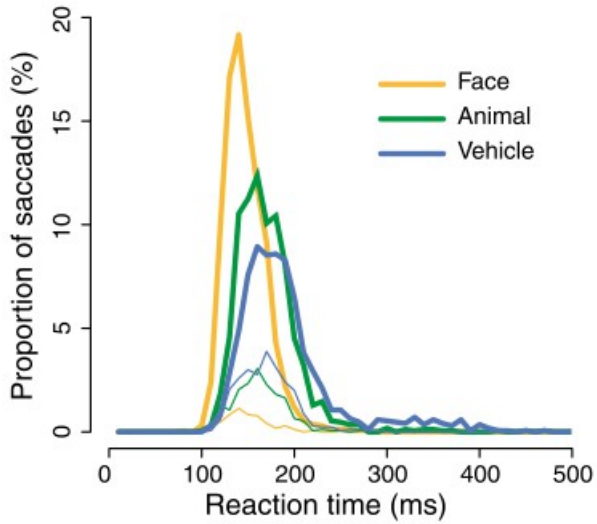
- Stimulus sampling bias

  ? Representation of agentic action plays a dominant role
- Inadequate computational methods for data analysis and modeling

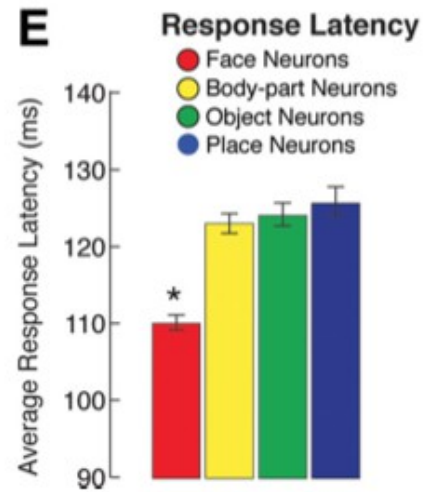
  ? Modeling functional architecture as a high-dimensional space captures complexity better than category-specific regions

Ultrarapid face detection
(Crouzet, Kirchner, Thorpe, 2010)

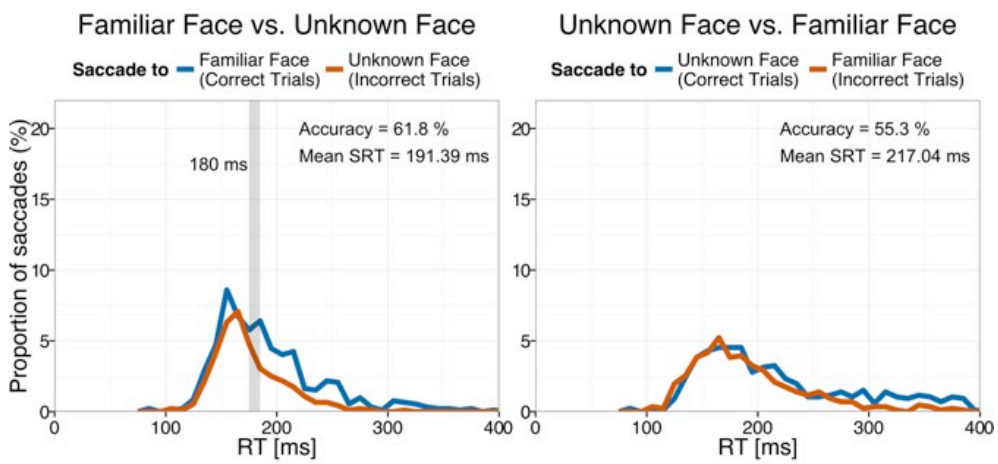




Face neurons are faster
(Bell et al. 2011)



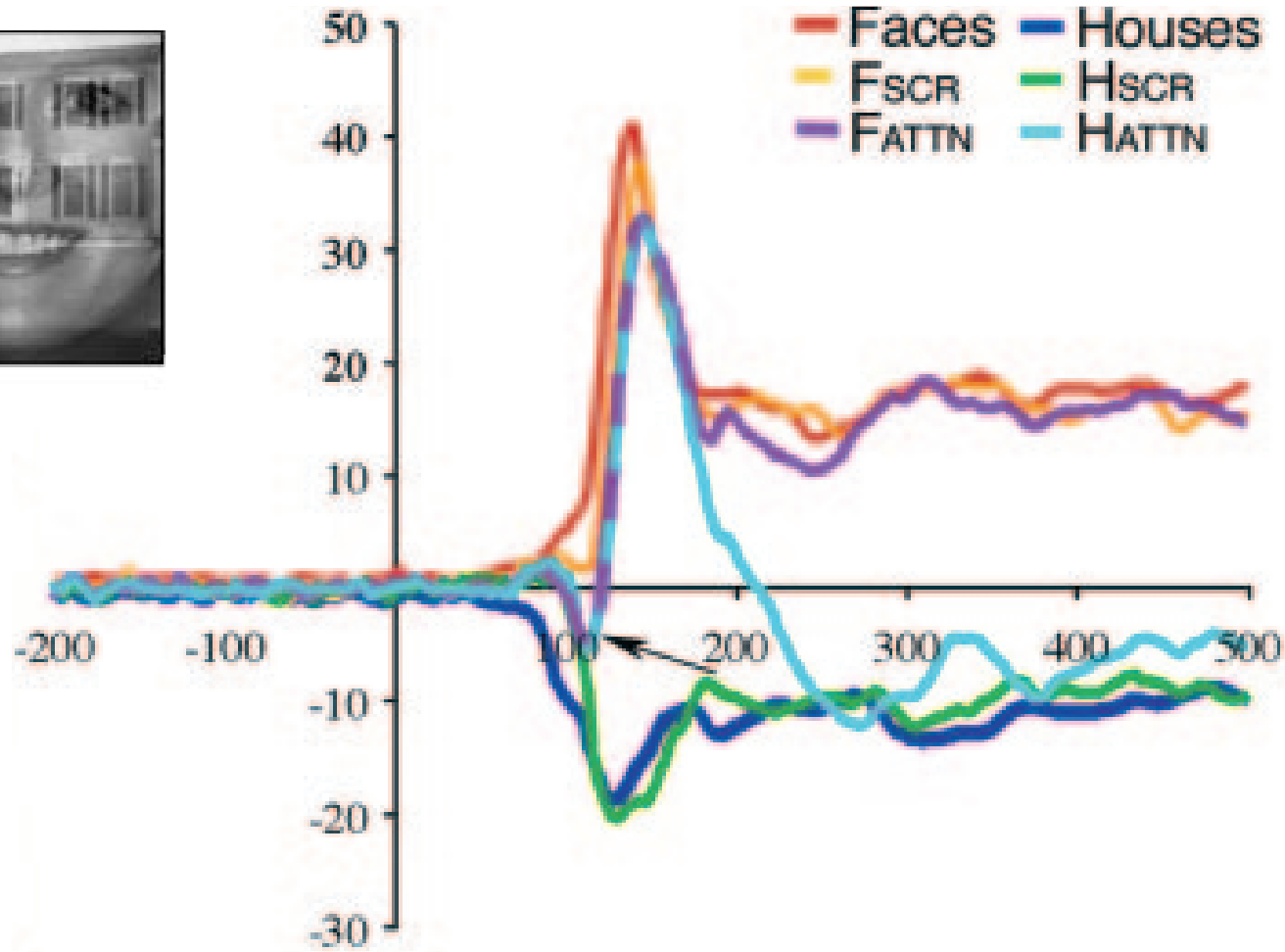
Very rapid detection of familiar faces
(Visconti di Oleggio Castello, Gobbini, et al. 2015)



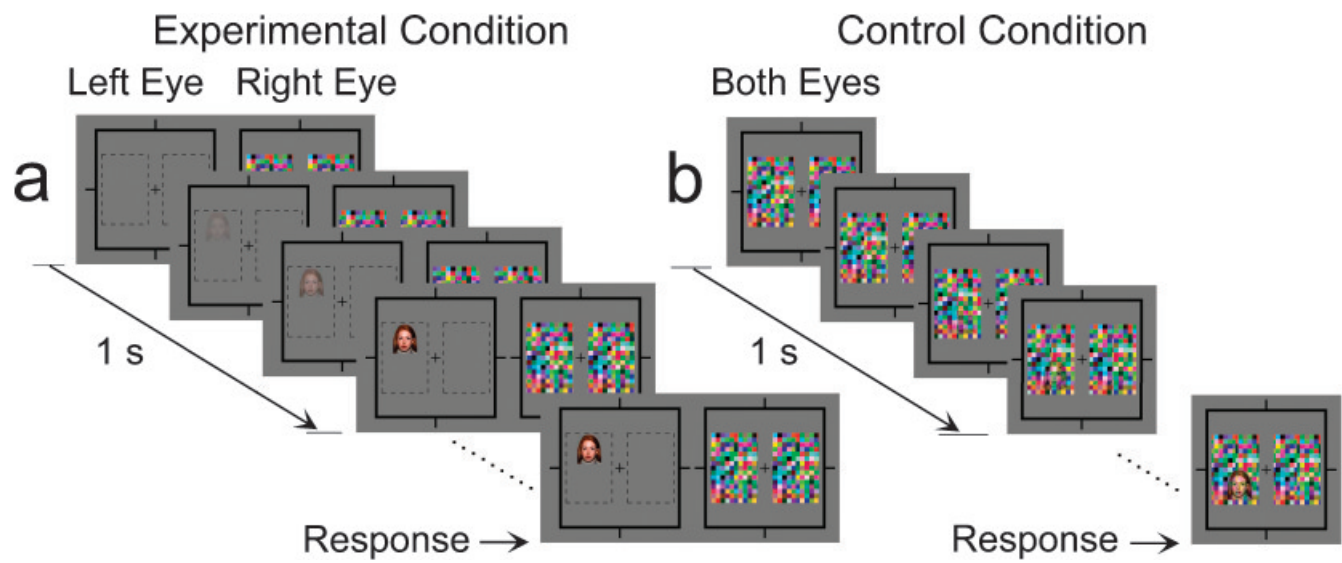
Early face-selective neural responses are not affected by attention

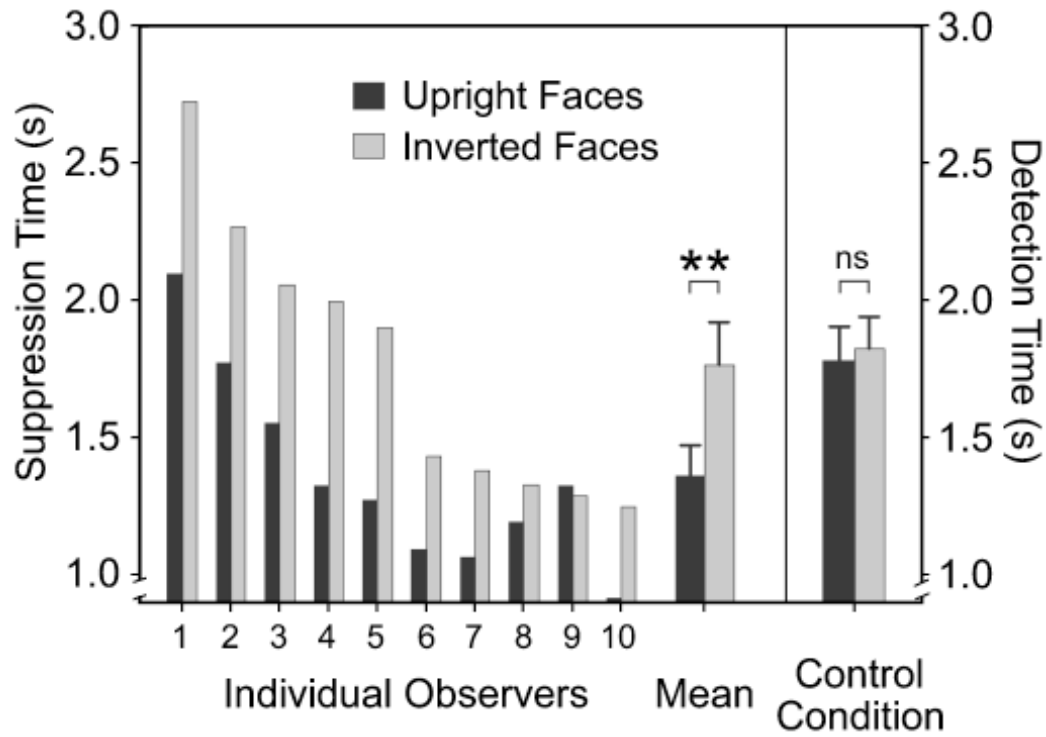
(Furey et al., 2006)

Selective attention to faces (FATTN) or houses (HATTN)

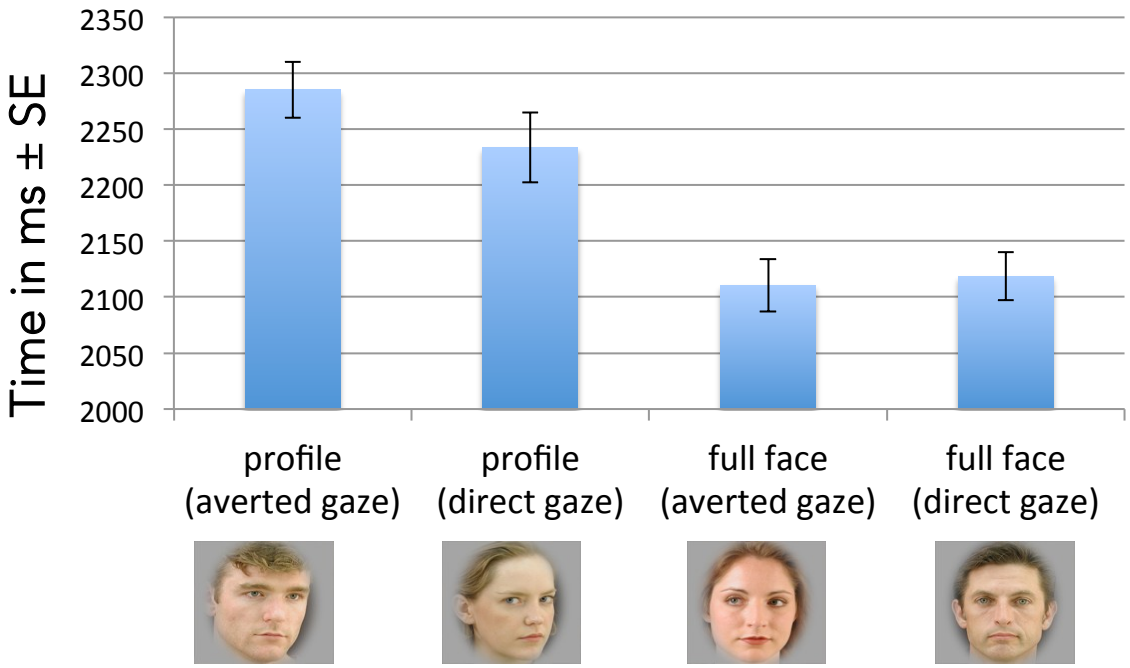


Upright faces break through interocular suppression faster than inverted faces
(Jiang, Costello, He, 2007)

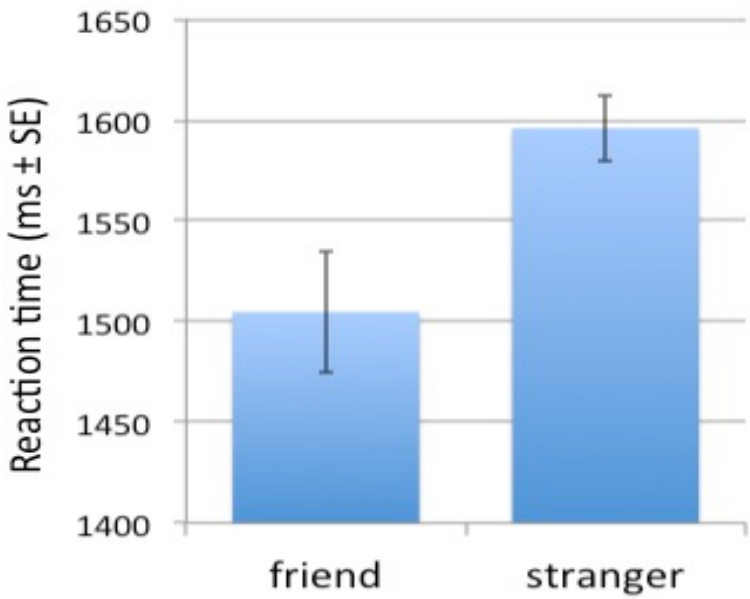




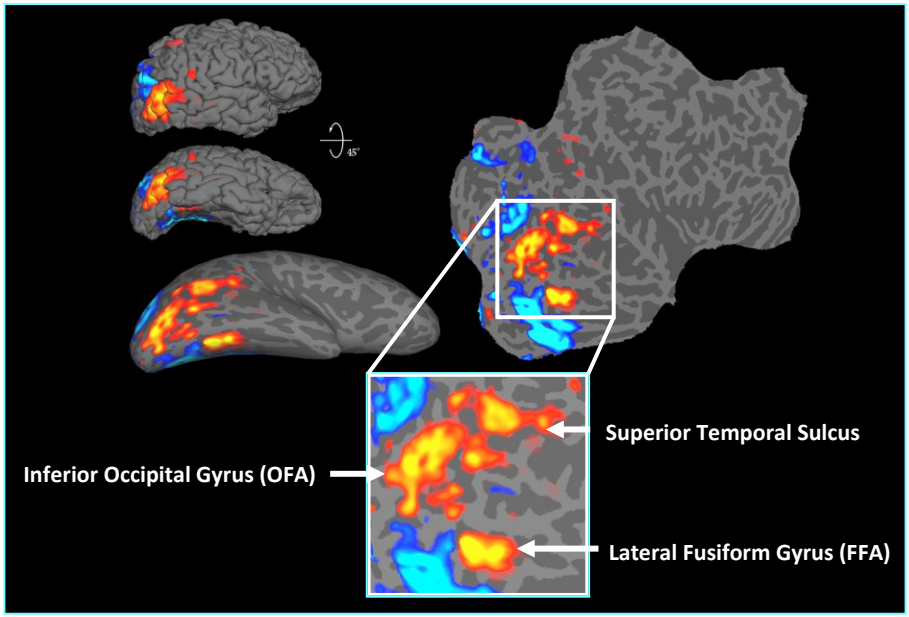
Faces in full view break through interocular suppression faster than faces in profile
(Gobbini et al., Consc Cogn, 2013)

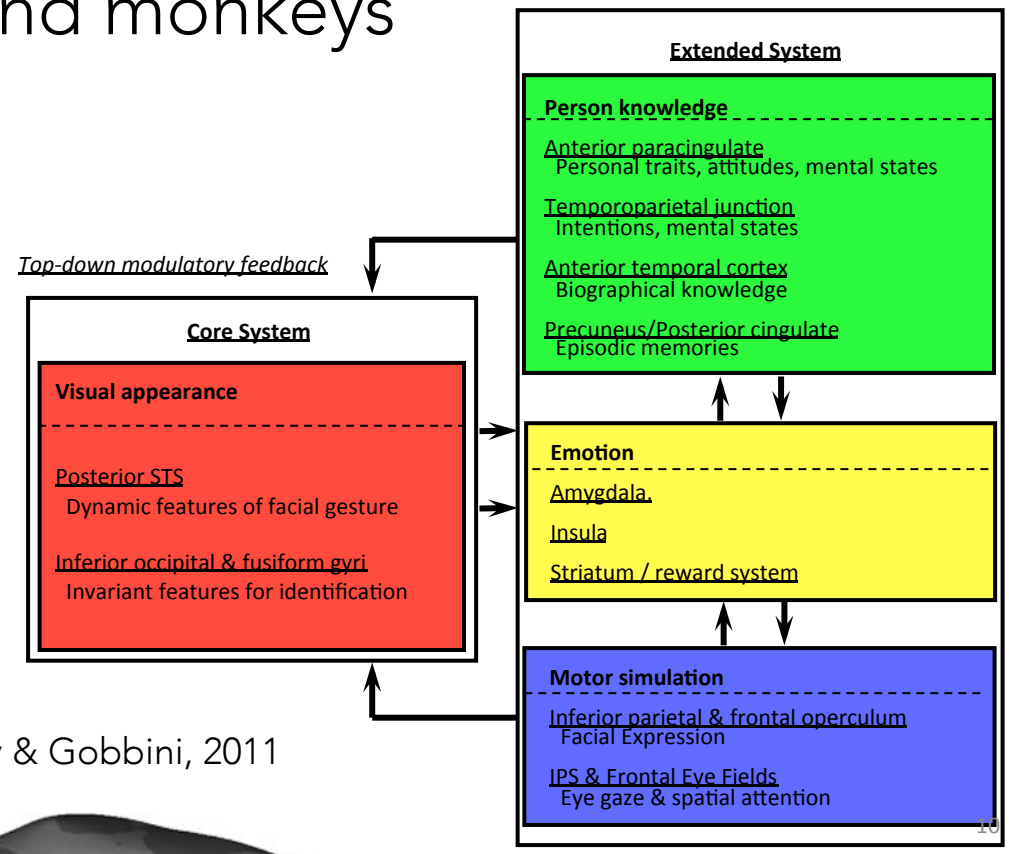


…and personally familiar faces break through faster than faces of strangers
(Gobbini et al. PLoS One, 2013)

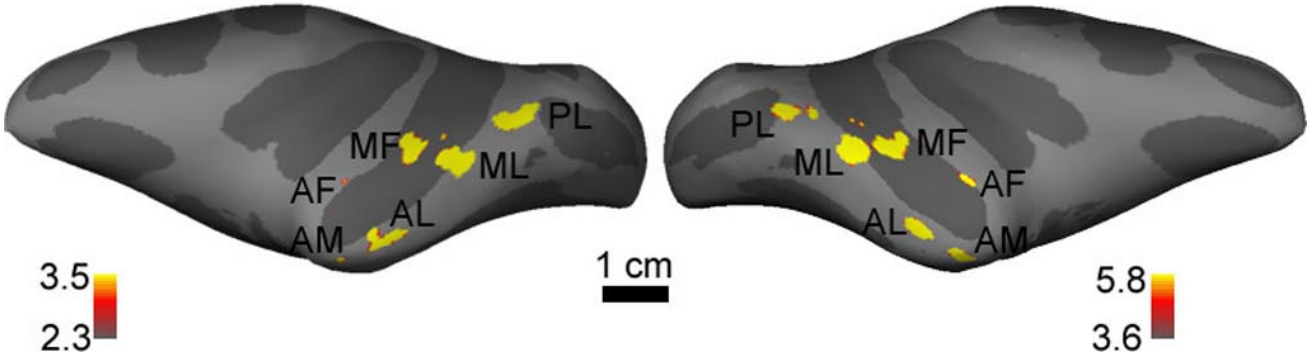


Distributed neural system for face perception in humans and monkeys





Haxby et al. 2000; Gobbini & Haxby, 2007; Haxby & Gobbini, 2011



Tsao et al. 2008

Neural representation of faces, bodies, and objects in ventral temporal cortex?

Optimized for person perception

- Face processing is fast and requires minimal attention and awareness
- Mediated by a distributed neural system in humans and monkeys

A fly in the ointment

- Stimulus sampling bias

  ? Representation of agentic action plays a dominant role
- Inadequate computational methods for data analysis and modeling

  ? Modeling functional architecture as a high-dimensional space captures complexity better than category-specific regions

A fly in the ointment

A fly in the ointment

Neural representation of faces, bodies, and objects in ventral temporal cortex?

Optimized for person perception

- Face processing is fast and requires minimal attention and awareness
- Mediated by a distributed neural system in humans and monkeys

A fly in the ointment

- Stimulus sampling bias

  ? Representation of agentic action plays a dominant role
- Inadequate computational methods for data analysis and modeling

  ? Modeling functional architecture as a high-dimensional space captures complexity better than category-specific regions

What's wrong with current models of representation of faces, bodies, and objects in ventral temporal cortex?

- Serious stimulus sampling bias
  - Still images
  - Limited range of categories

- Inadequate computational methods for data analysis and modeling

What's wrong with current models of representation of faces, bodies, and objects in ventral temporal cortex?

- Serious stimulus sampling bias
  - Still images
  - Limited range of categories

- Inadequate computational methods for data analysis and modeling
  - Univariate contrasts
  - Wrong question: What is the function of an area?

What's wrong with current models of representation of faces, bodies, and objects in ventral temporal cortex?

- Serious stimulus sampling bias
  - Still images
  - Limited range of categories

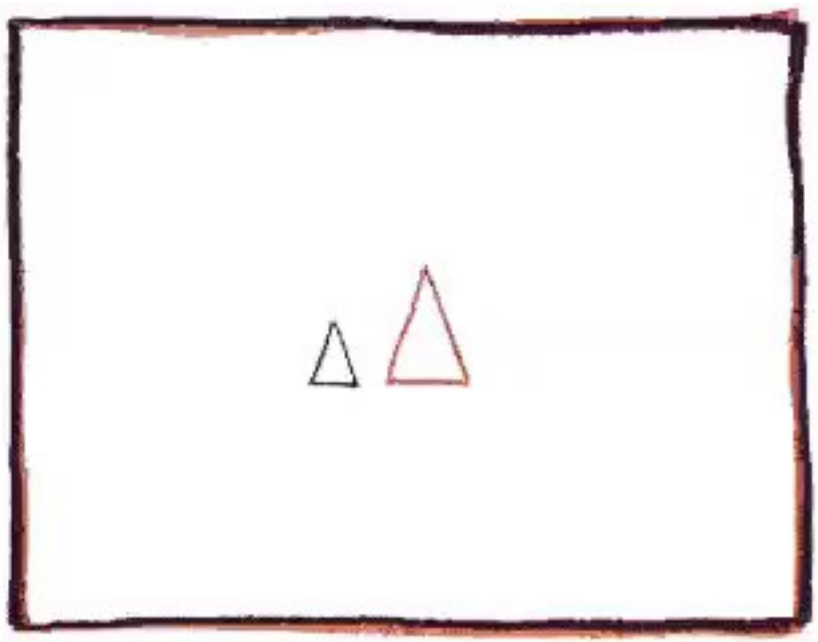
- Inadequate computational methods for data analysis and modeling
  - Univariate contrasts
  - Wrong question: What is *the function* of an area?

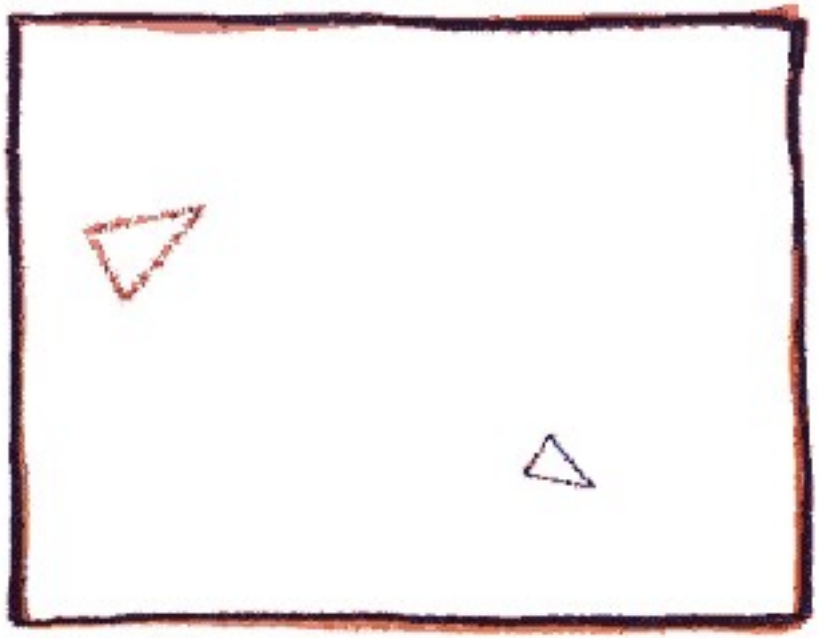
What's wrong with current models of representation of faces, bodies, and objects in ventral temporal cortex?

- Serious stimulus sampling bias
  - Still images
  - Limited range of categories

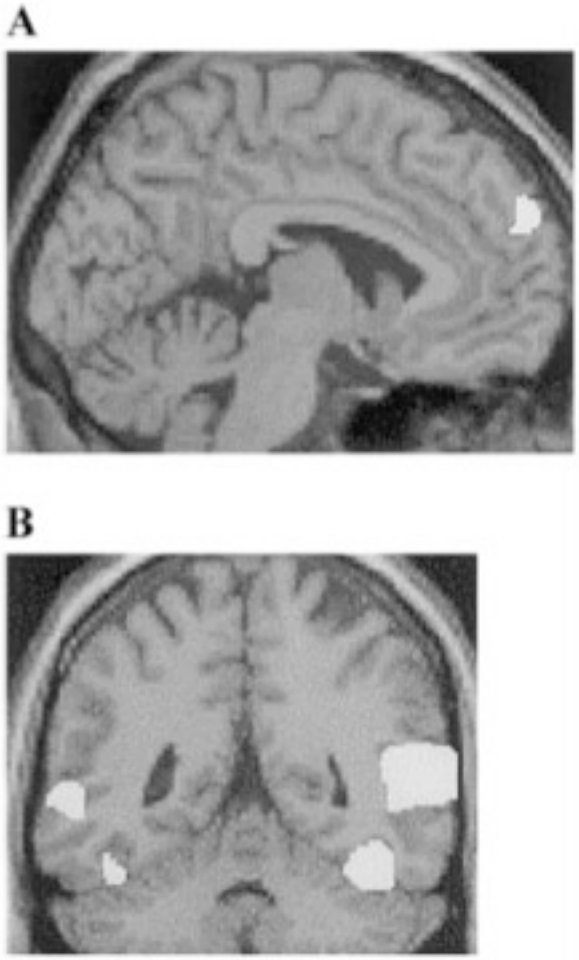
- Inadequate computational methods for data analysis and modeling
  - Univariate contrasts
  - Wrong question: What is *the function* of an *area*?

Social Animations
(Heider-Simmel)



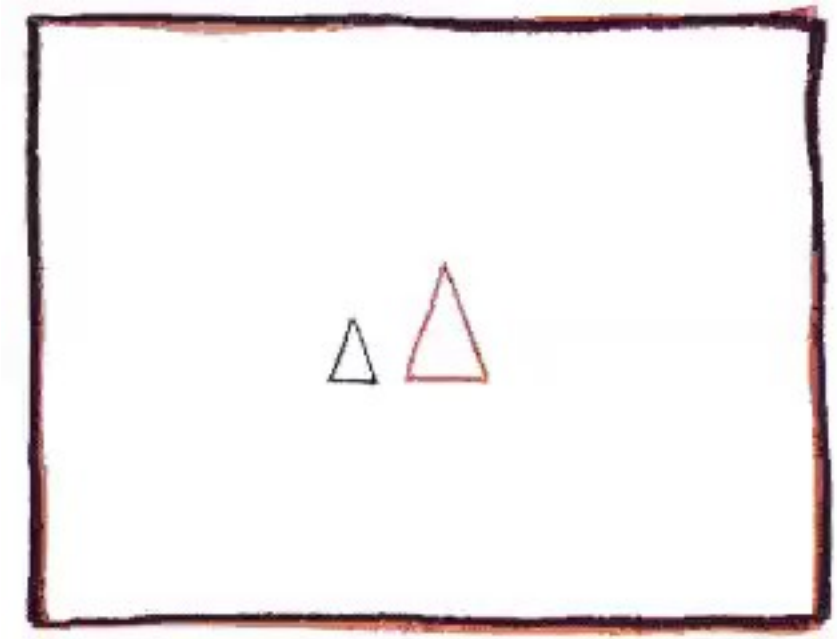


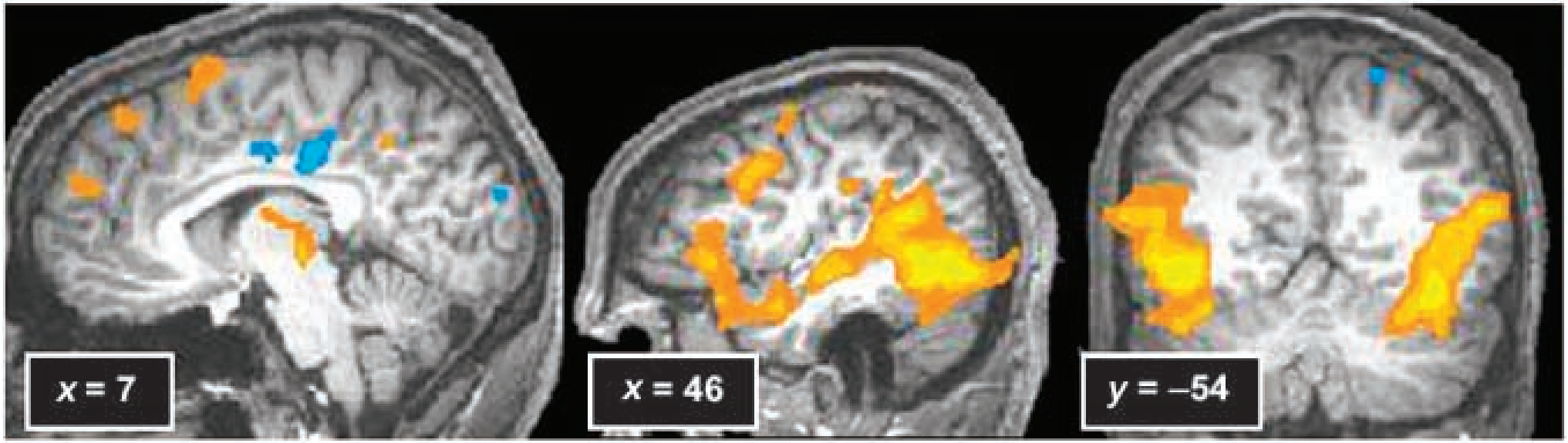
Social Animations
(Castelli et al. 2000)



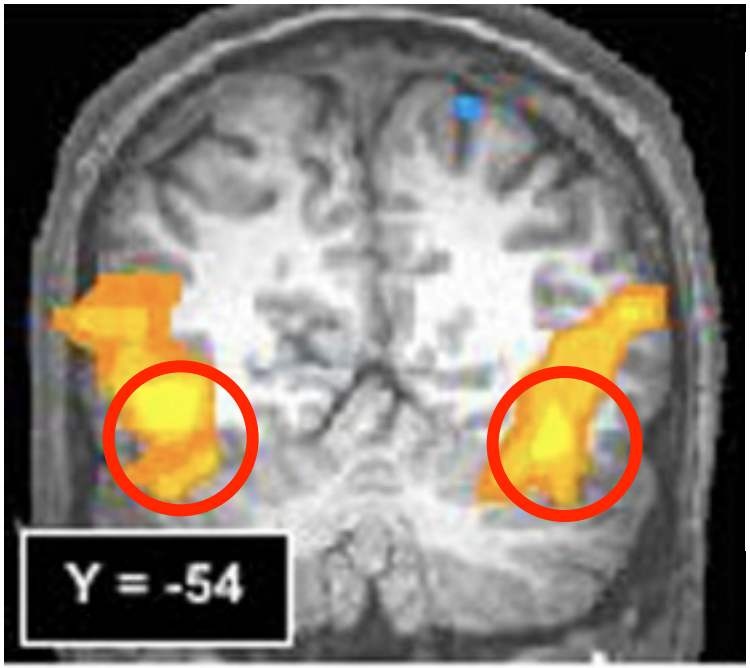
Social vs random animations

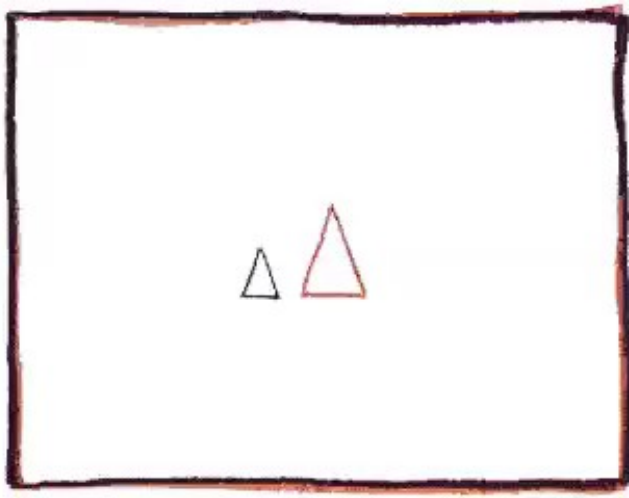
Extensive activation of perceptual and social brain systems by social animations
(Gobbini et al., 2007)





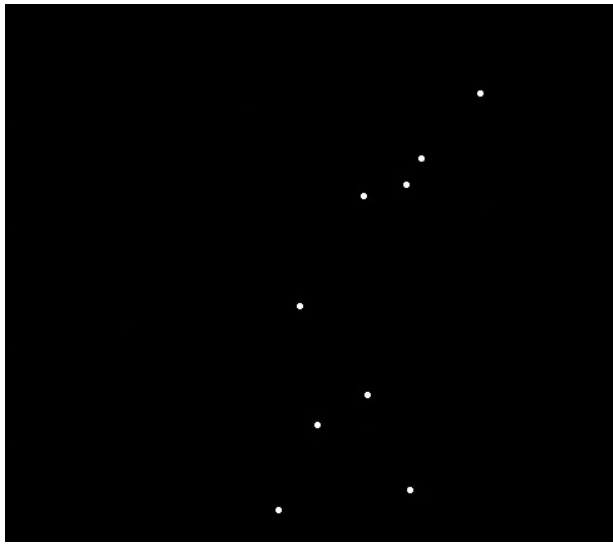
Social animations





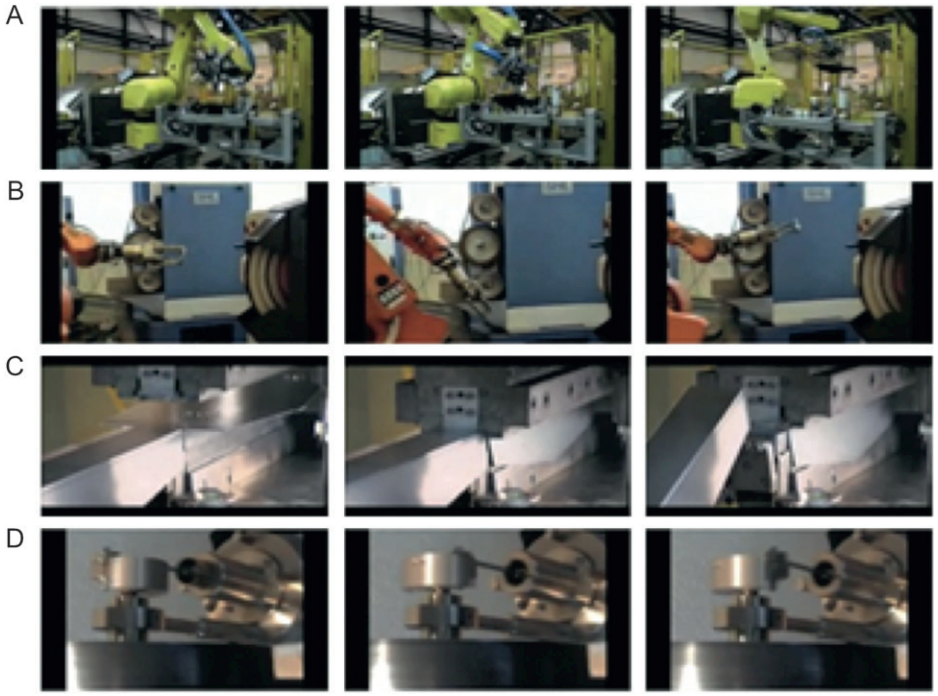
Fusiform activations without faces or bodies
(Gobbini et al. 2007)
(see also Schultz et al. 2003)

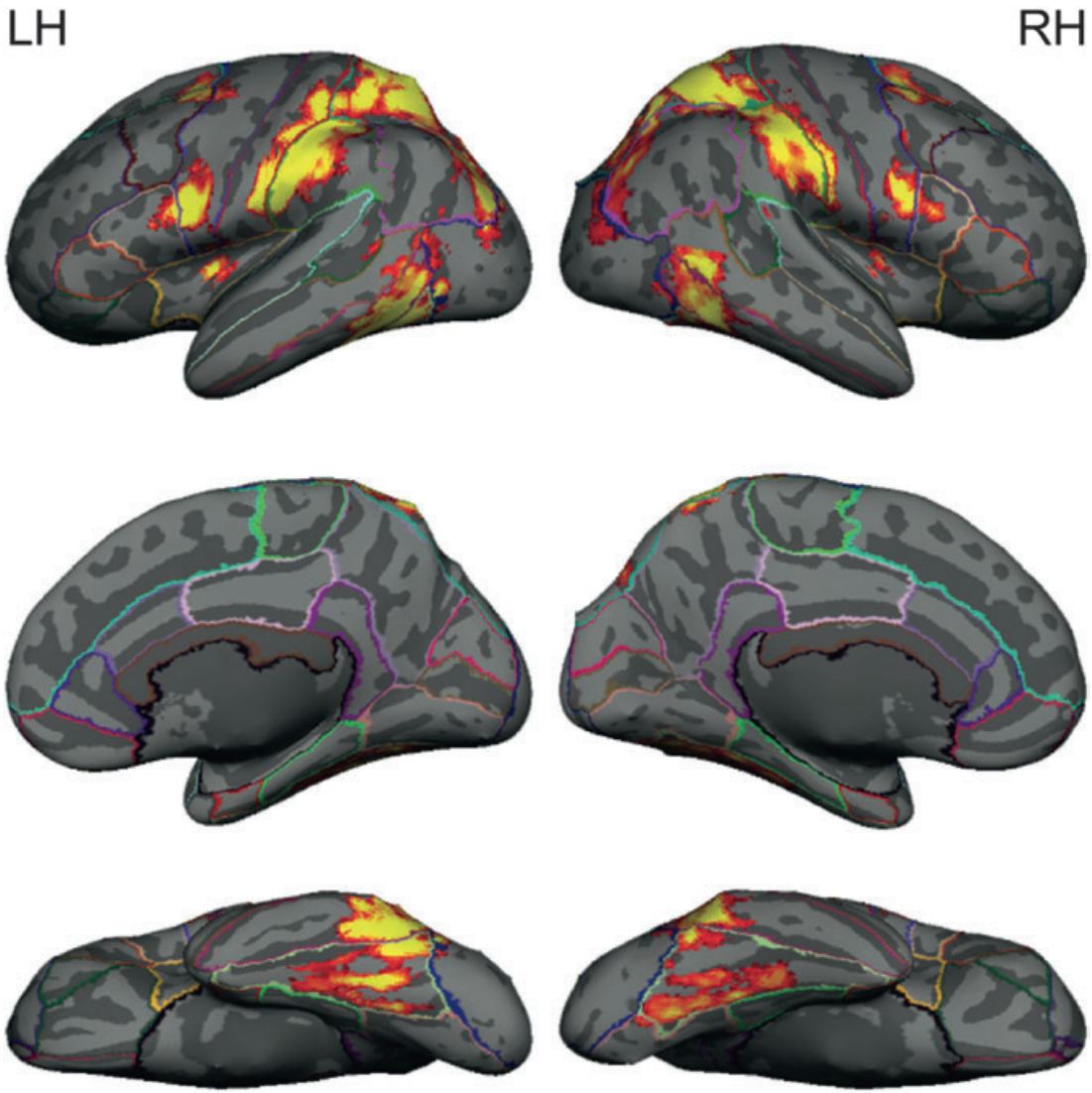
Biological motion
(point-light displays)



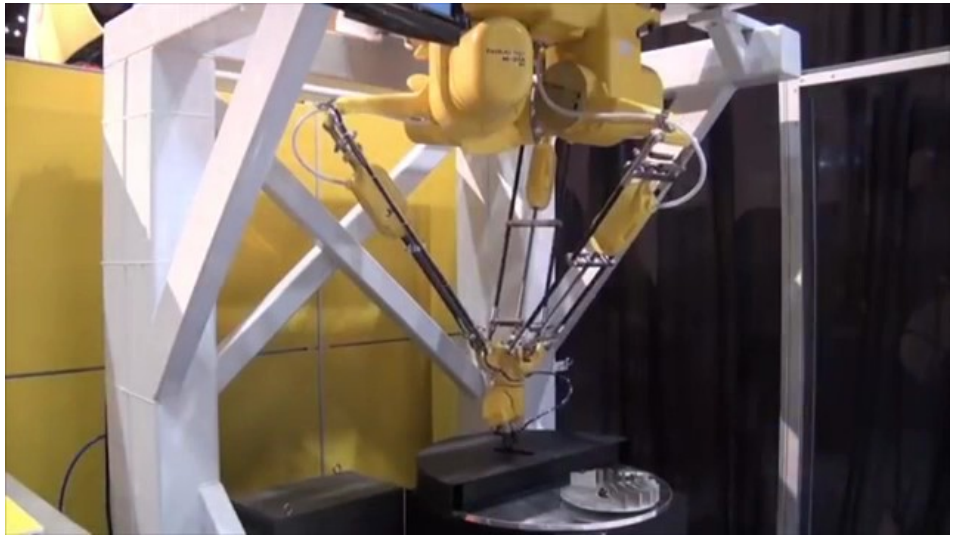
Goal-directed actions performed by non-human-like robots activate the FFA

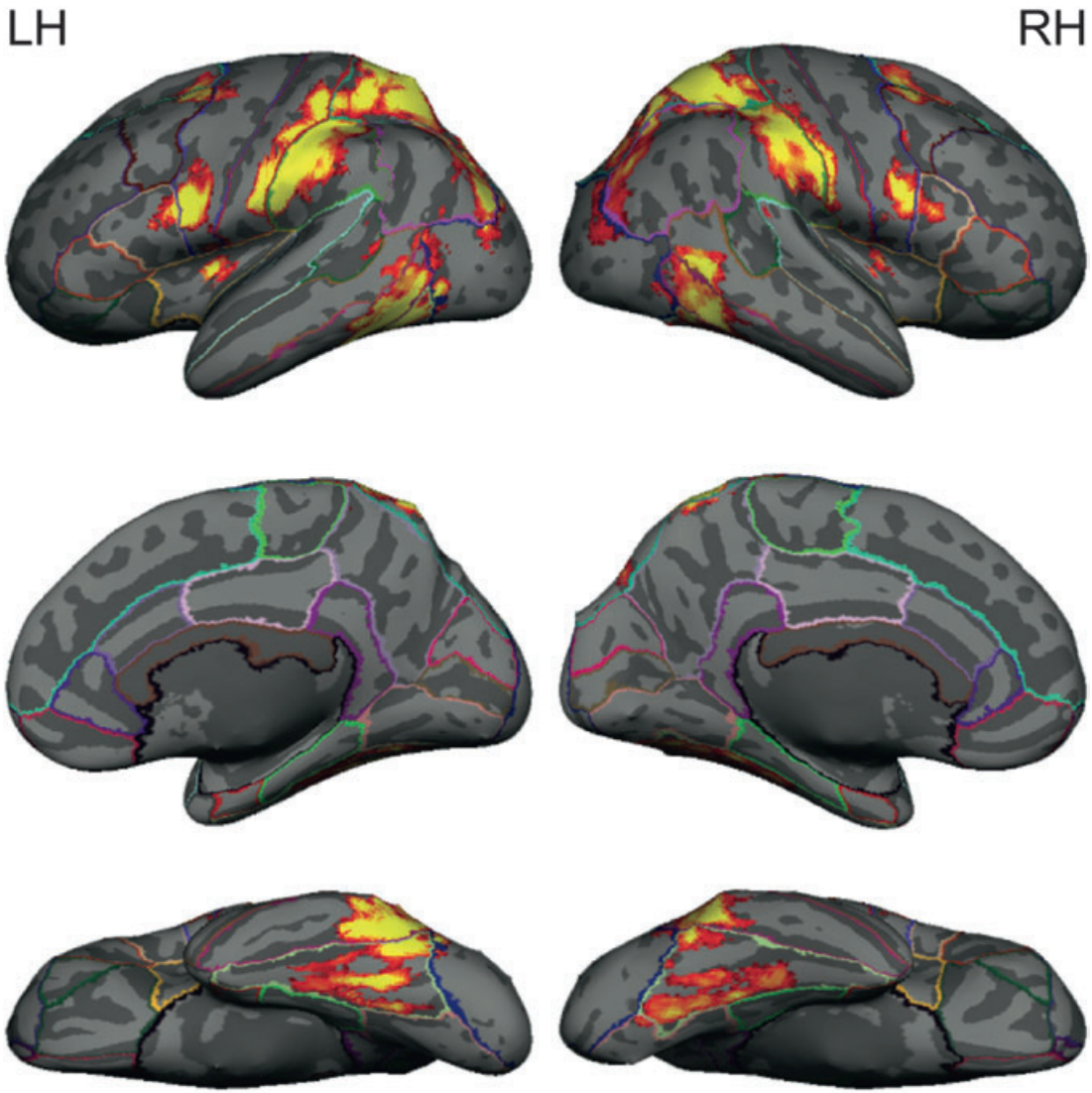
(Shultz & McCarthy, 2012)





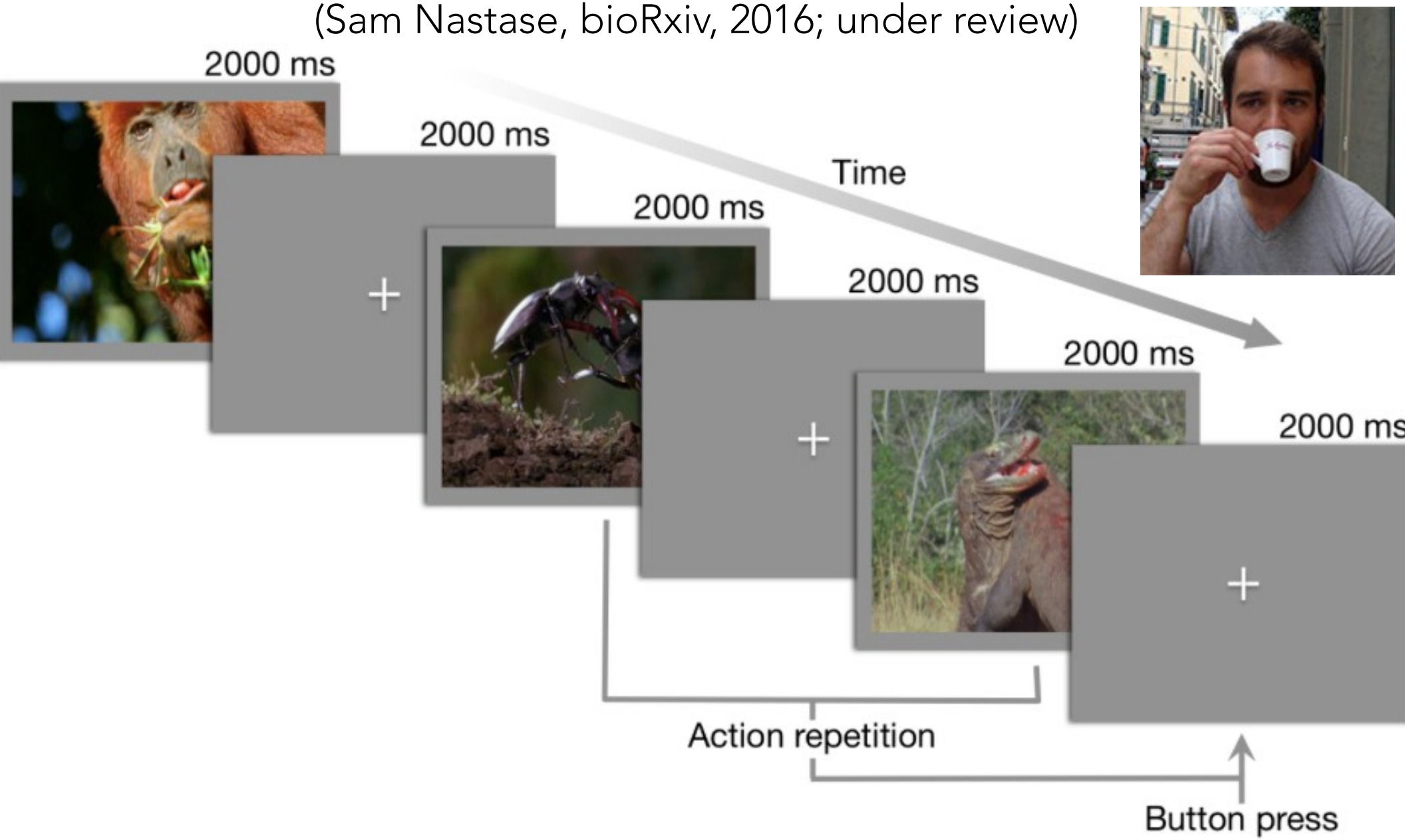
Goal-directed actions performed by non-human-like robots activate the FFA
(Shultz & McCarthy, 2012)

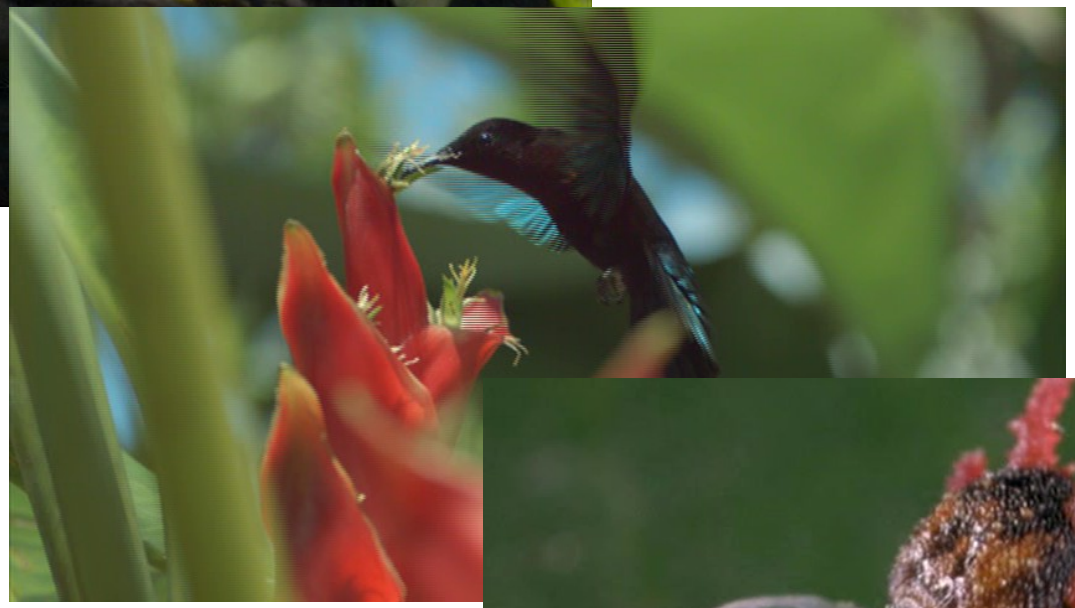


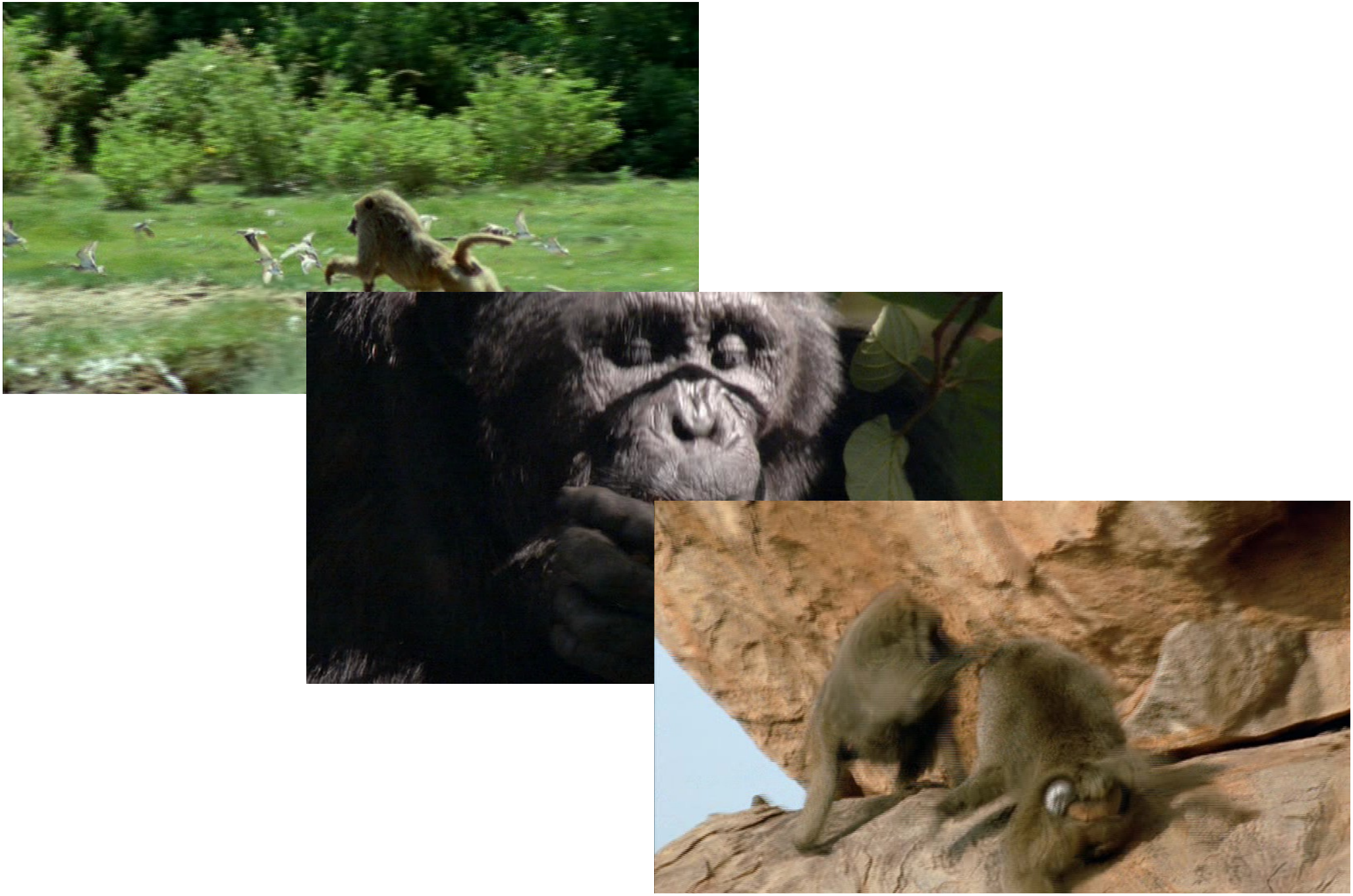


Representation of action in the cortex for perception of faces and animate stimuli

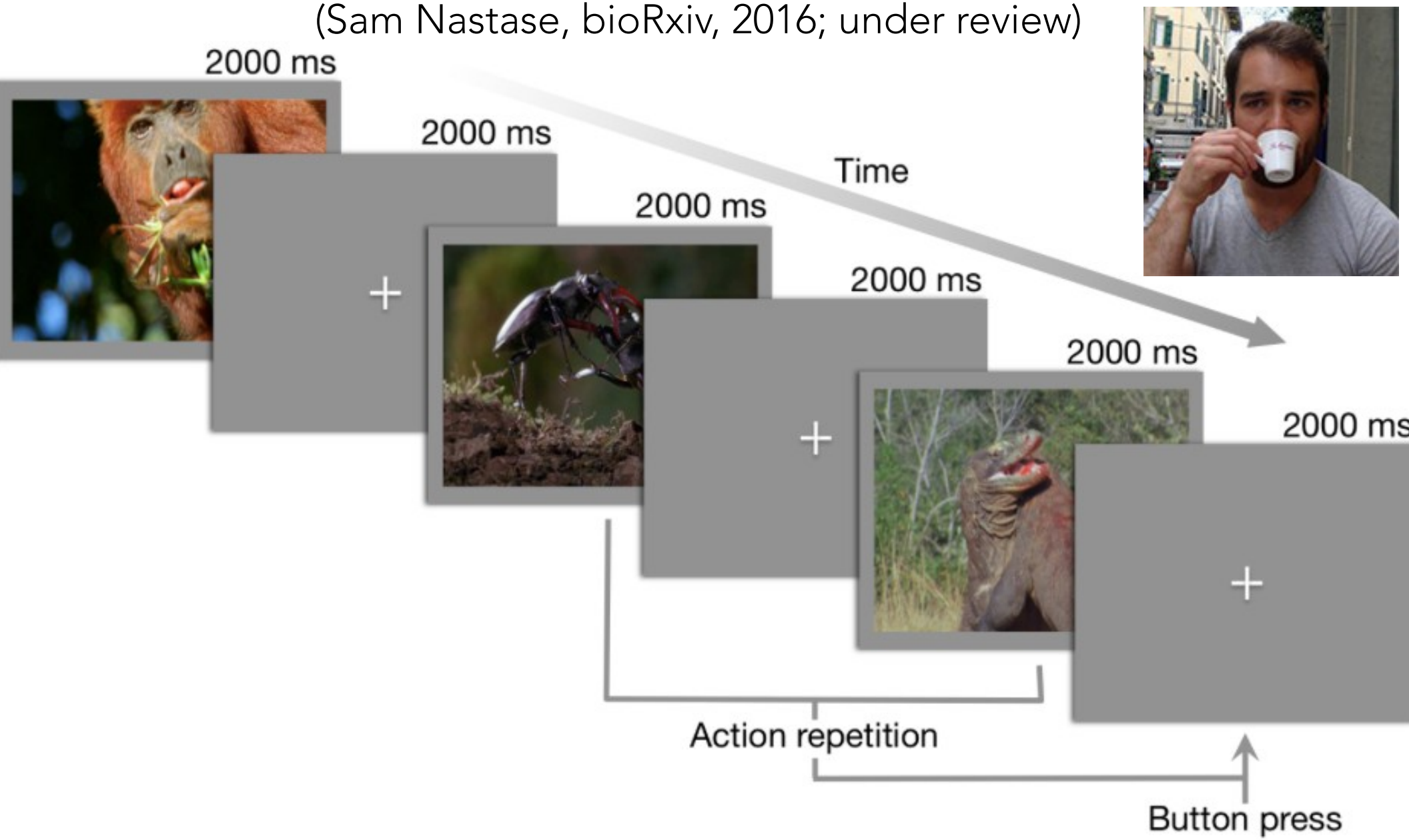
Study of the effect of attention on the
geometry of representations of animal taxa and animal behaviors
(Sam Nastase, bioRxiv, 2016; under review)



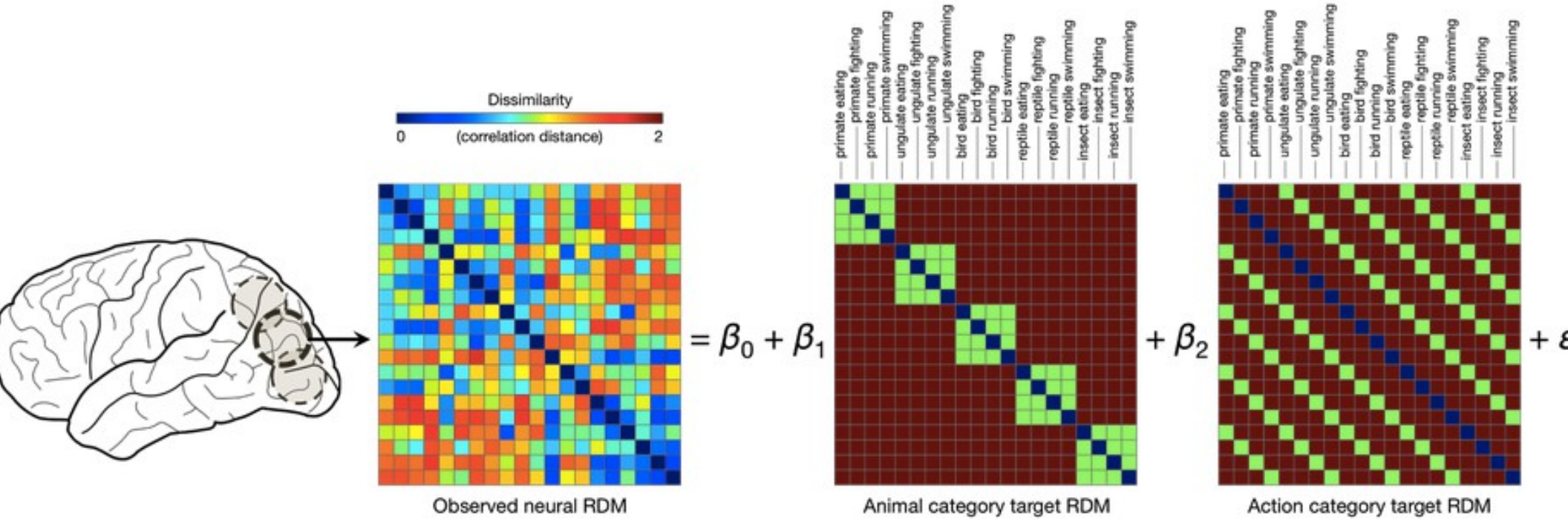




Study of the effect of attention on the
geometry of representations of animal taxa and animal behaviors
(Sam Nastase, bioRxiv, 2016; under review)

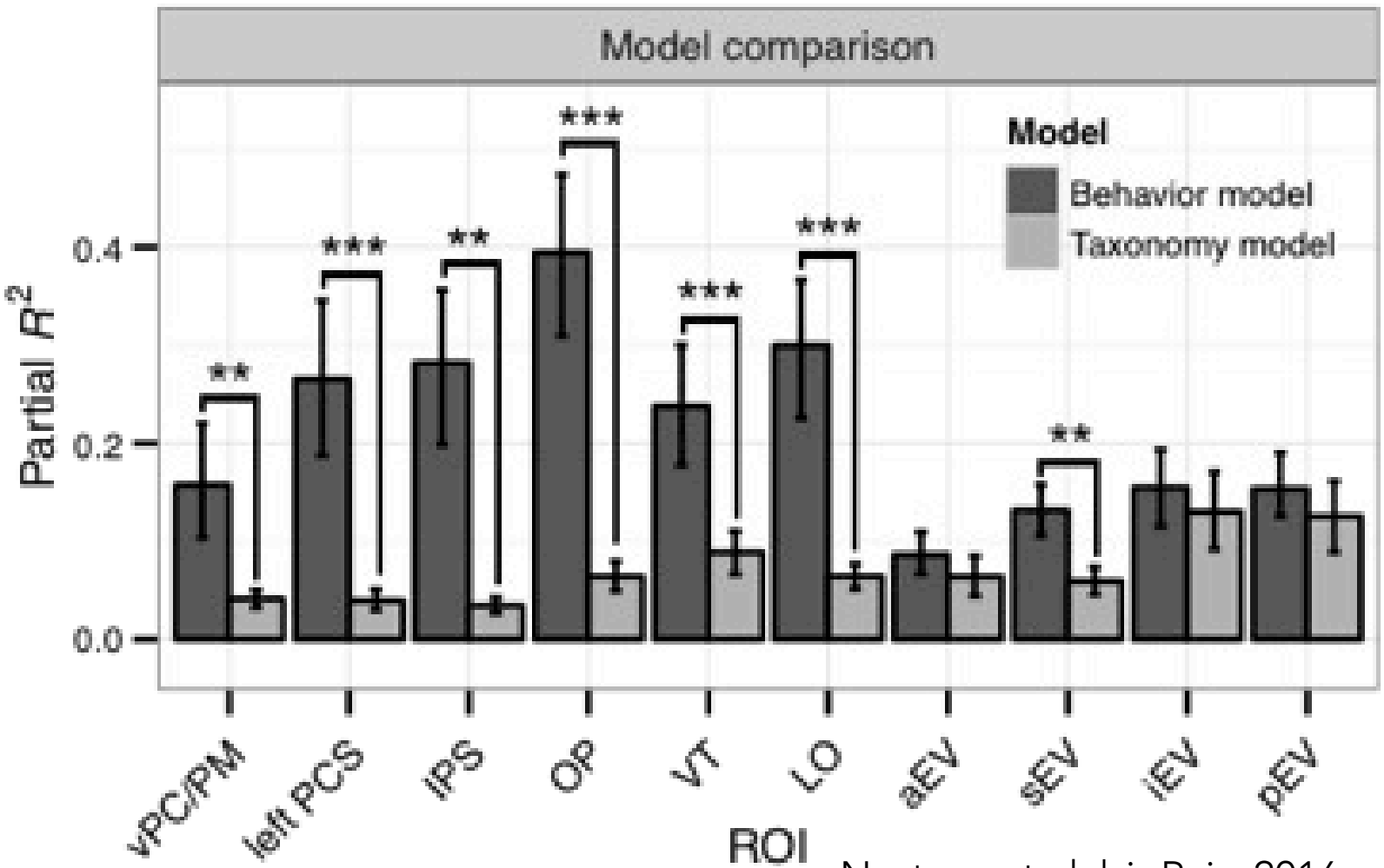


Regional representational geometry was modeled as a weighted sum of RDMs for animal taxa and action categories



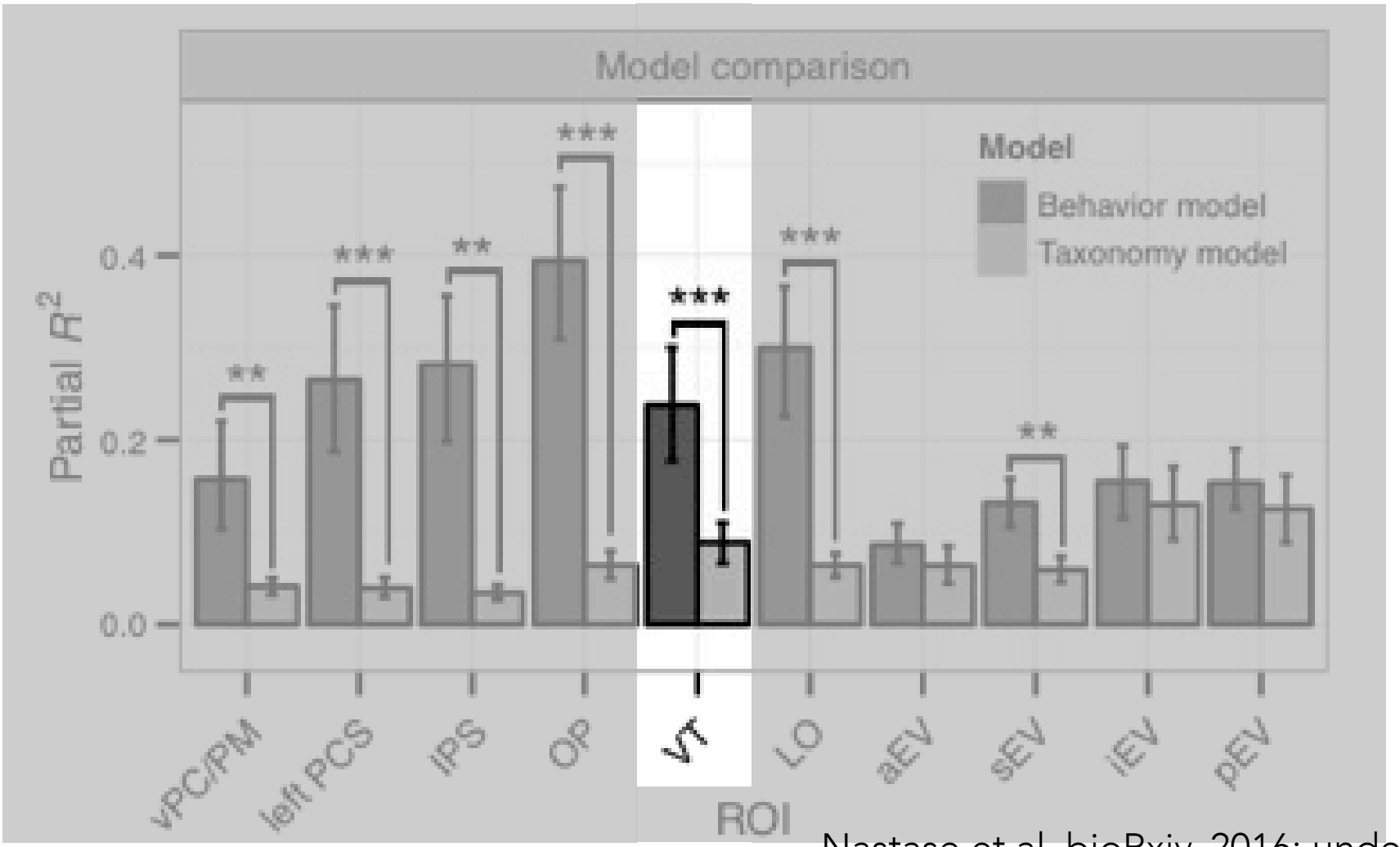
Nastase et al. bioRxiv, 2016; under review

Representational geometry is dominated by perceived behavior/action, even in ventral temporal cortex



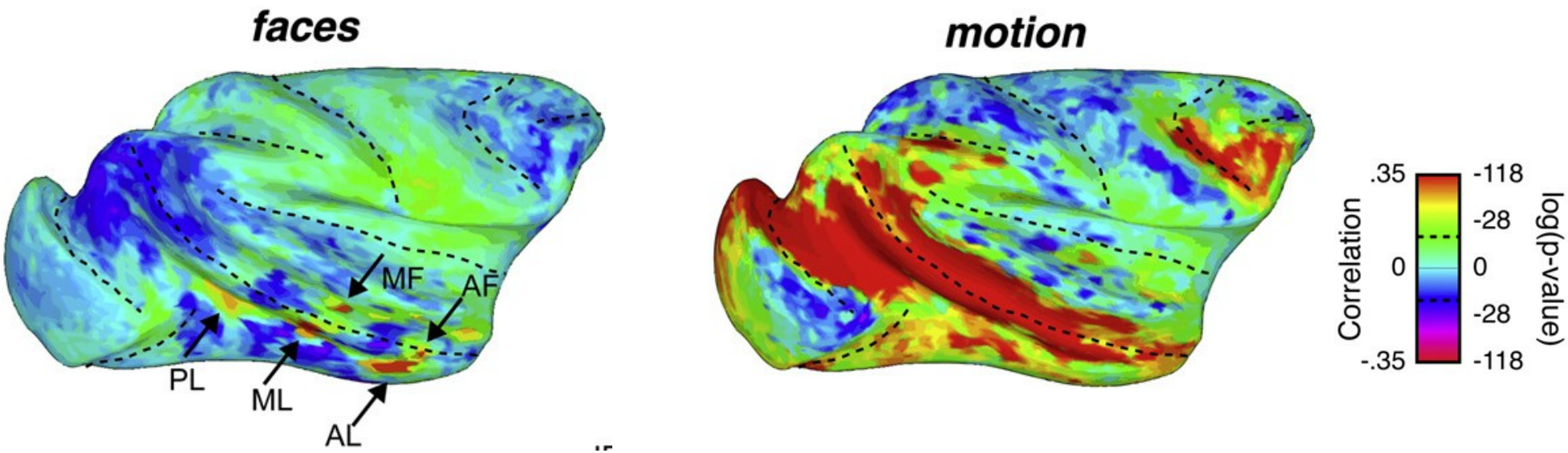
Nastase et al. bioRxiv, 2016; under review

Representational geometry is dominated by perceived behavior/action, even in ventral temporal cortex

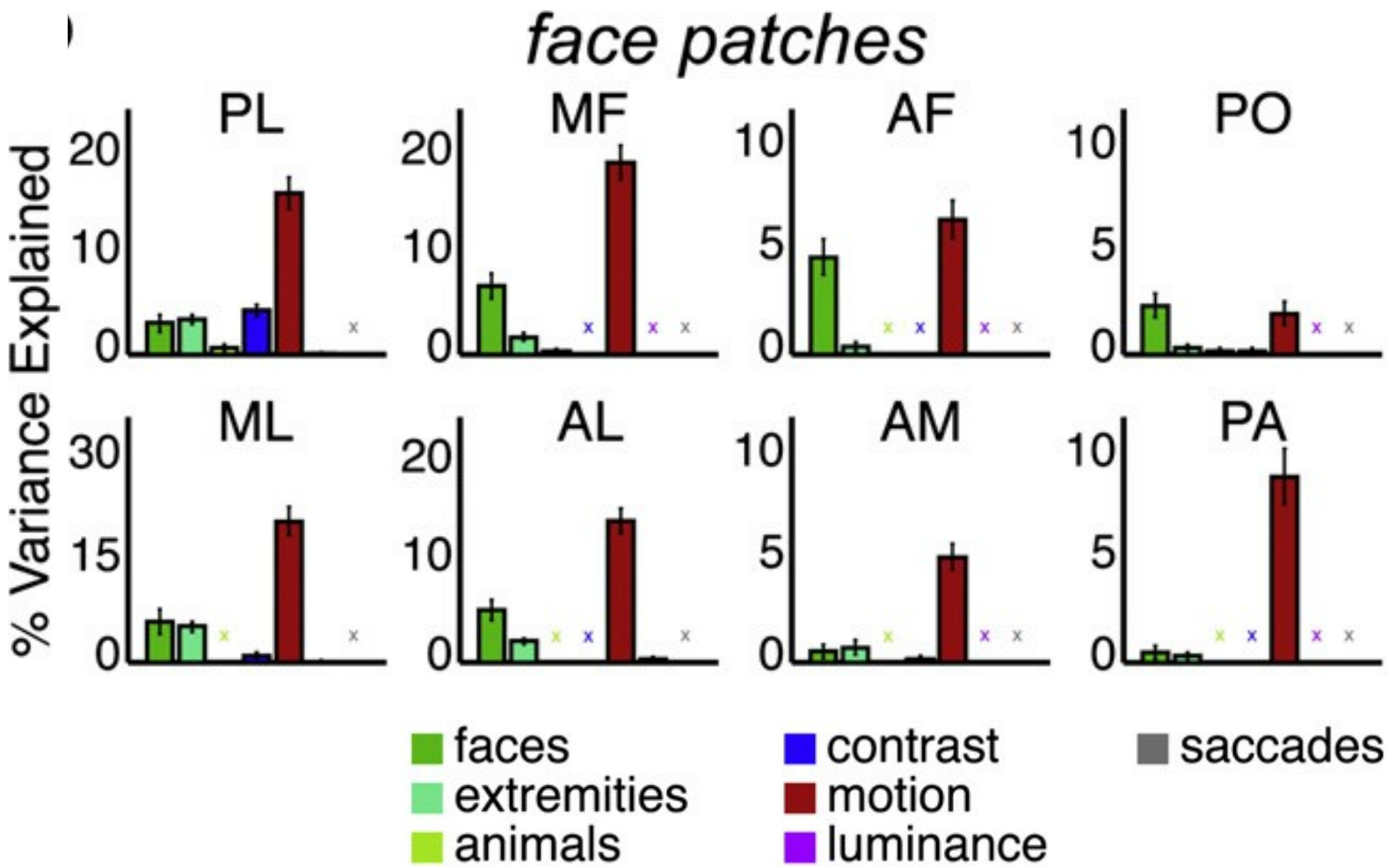


Nastase et al. bioRxiv, 2016; under review

Responses to dynamic natural movies in monkeys is dominated by motion (Russ & Leopold, 2015)

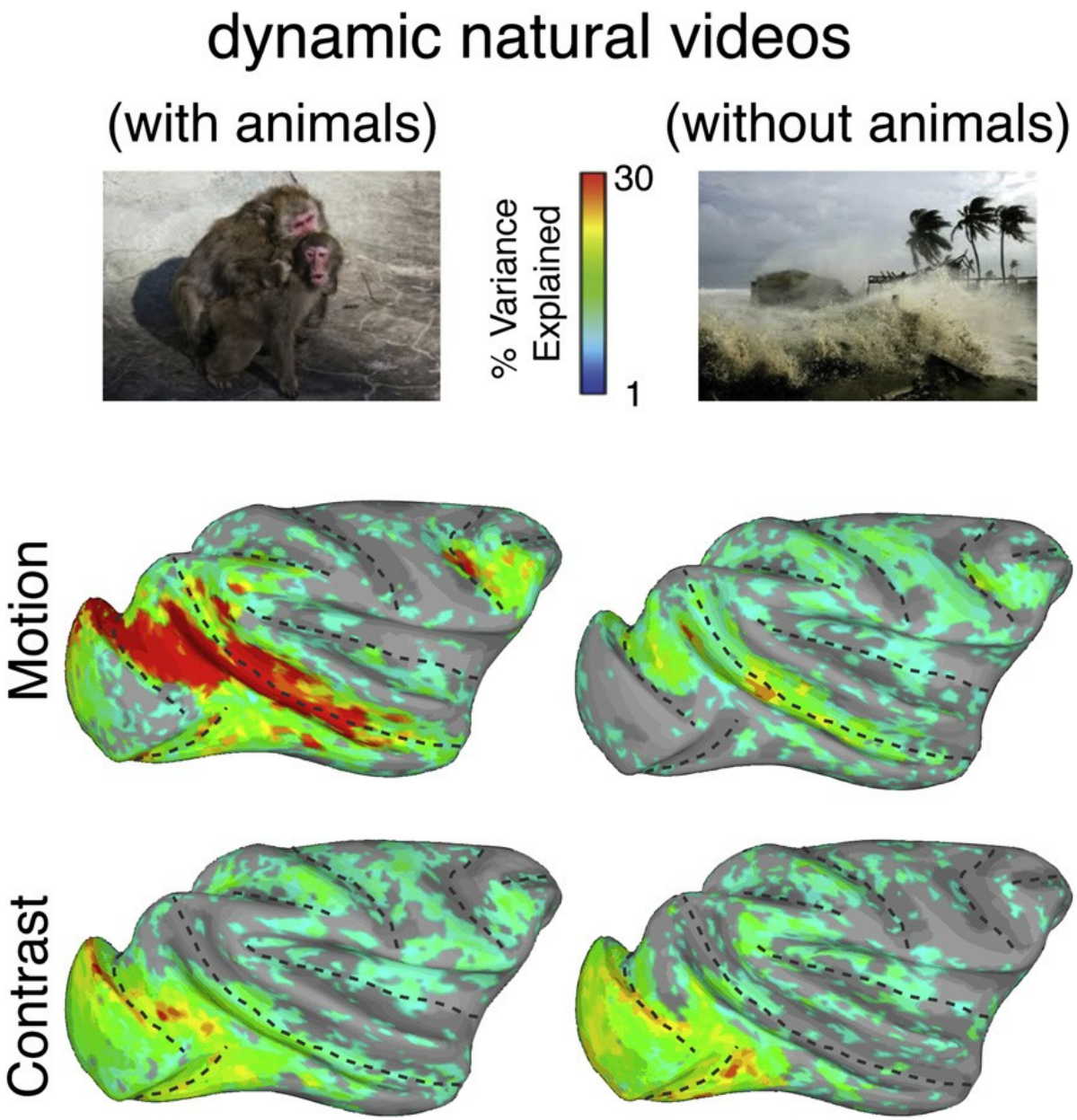


Responses to dynamic natural movies in *monkey face patches* is dominated by motion (Russ & Leopold, 2015)

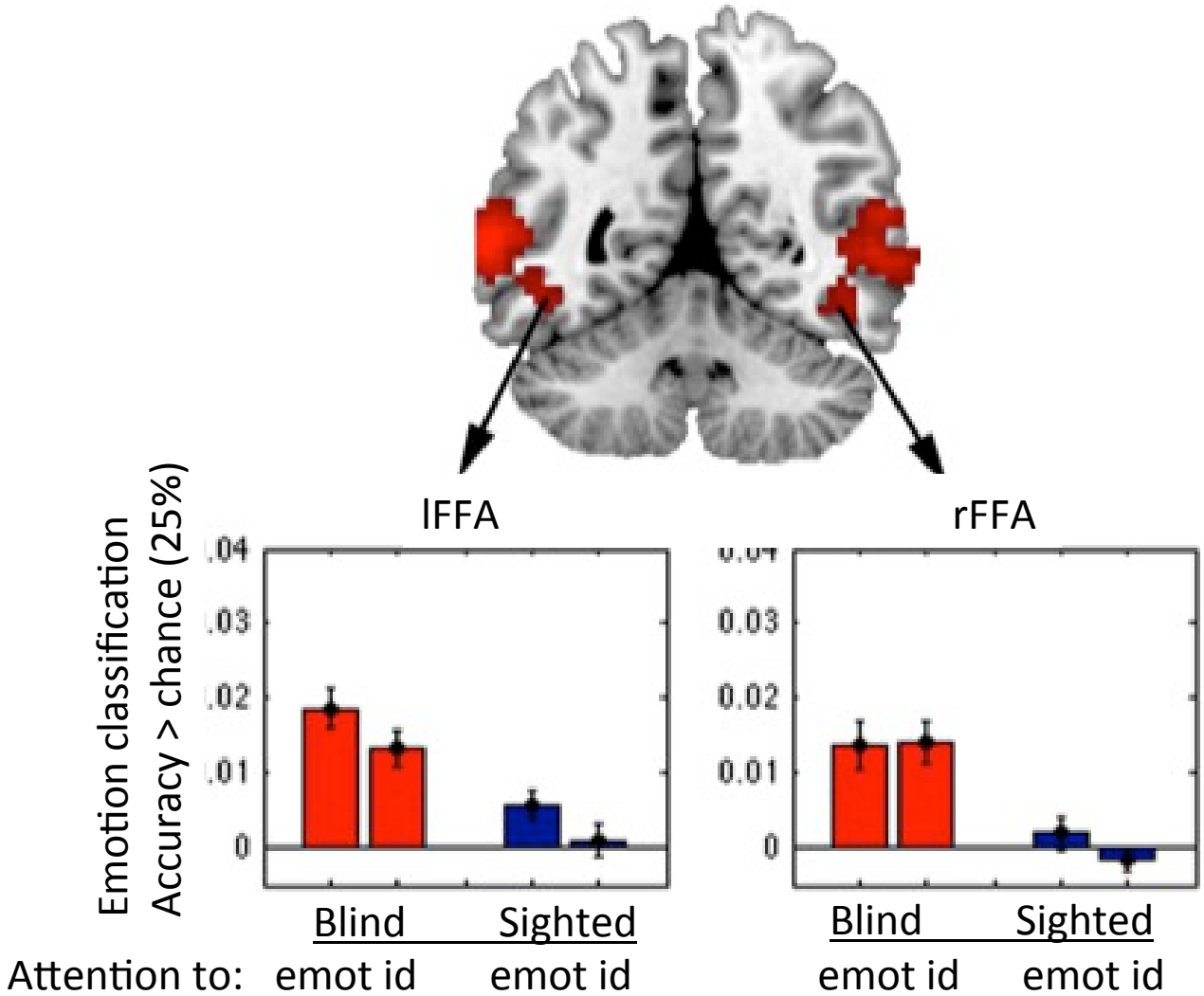


Responses to dynamic natural movies in monkeys is dominated by *animal motion*

(Russ & Leopold, 2015)



Lateral fusiform activity evoked by voice perception in the congenitally blind (Fairhall & Gobbini, under review)



Lateral fusiform gyrus activity is evoked or modulated by:

- Moving triangles
- Moving points of light
- Industrial robots
- Animal behavior more than animal form
- Voices in the congenitally blind

Stimulus sampling bias for *still images* of faces, bodies, and objects has restricted understanding of the functional architecture of ventral temporal cortex.

Lateral fusiform gyrus activity is evoked or modulated by:

- Moving triangles
- Moving points of light
- Industrial robots
- Animal behavior more than animal form
- Voices in the congenitally blind

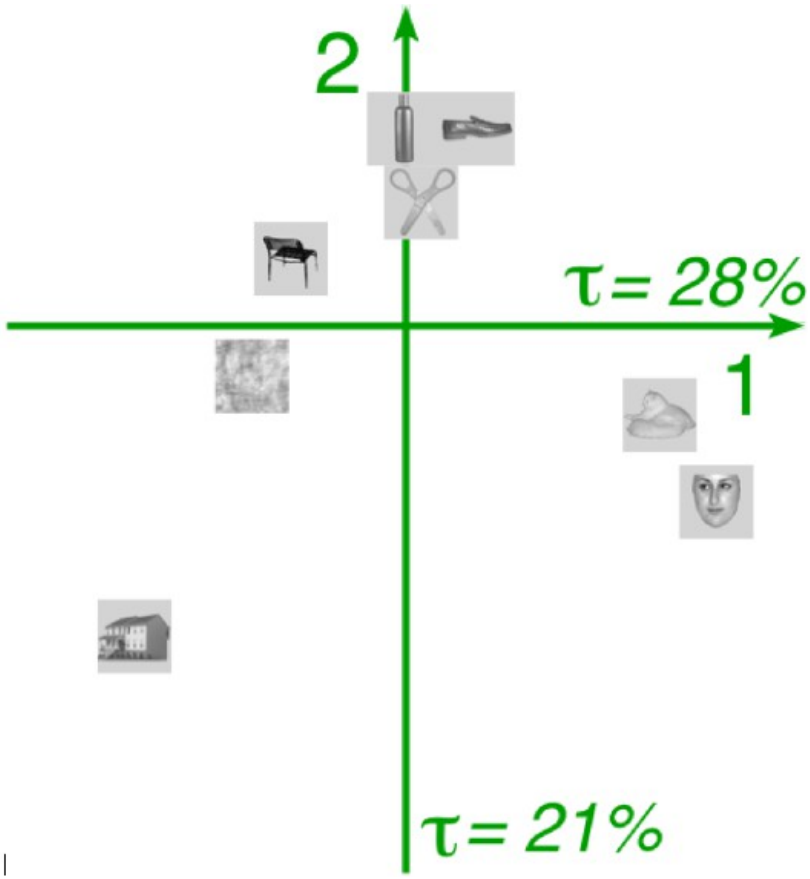
So, what is represented in the lateral fusiform gyrus?

- Animate as compared to inanimate entities?
  (Kriegeskorte, Mahon, Caramazza, Grill-Spector, and many more)
- Agency (Gobbini et al. 2007, 2011)

The representation of animacy in the ventral temporal lobe

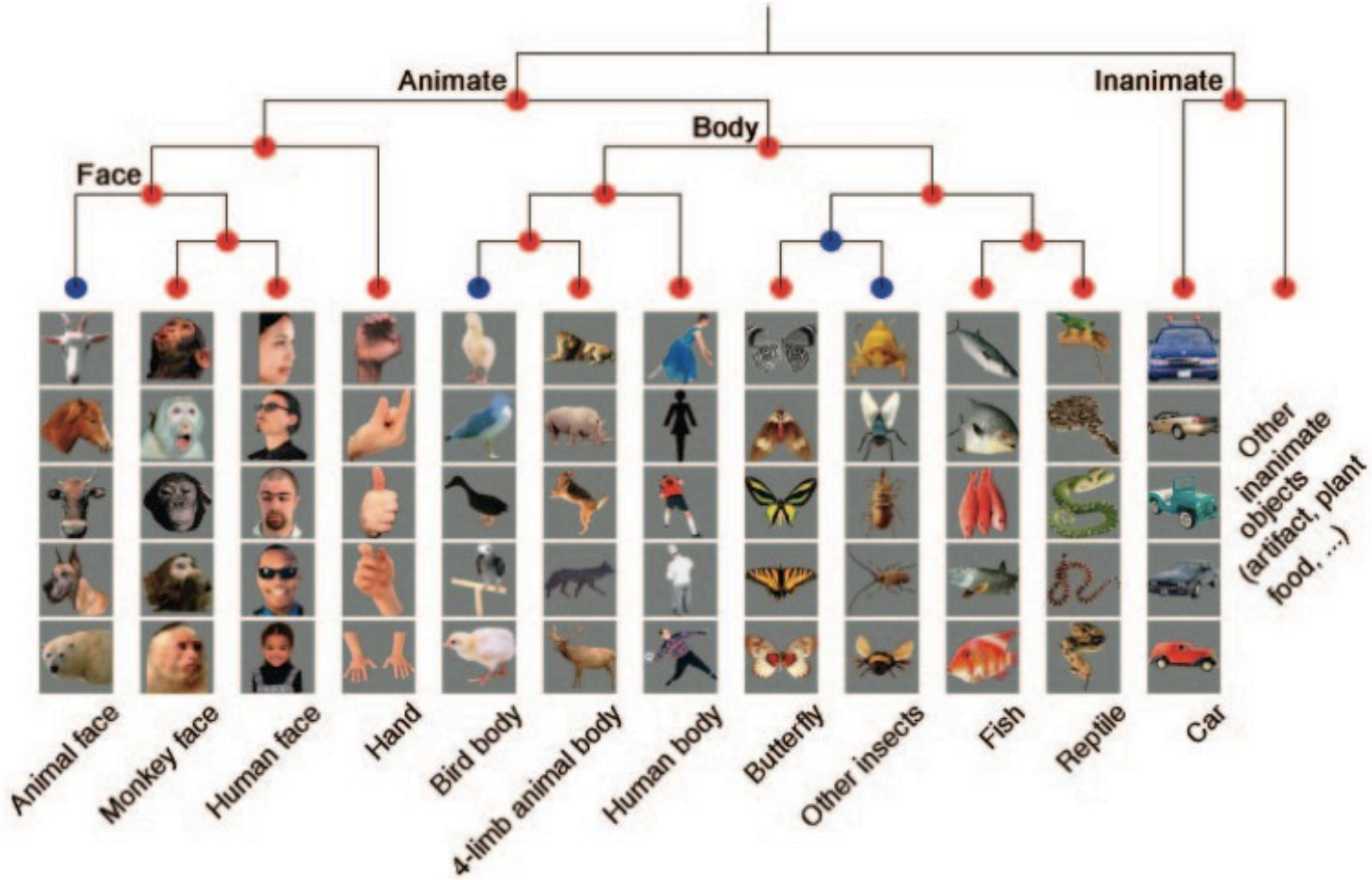
Is the animate-inanimate distinction a major large-scale feature?

Is the animate-inanimate distinction a major large-scale feature that distinguishes lateral from medial ventral temporal cortex



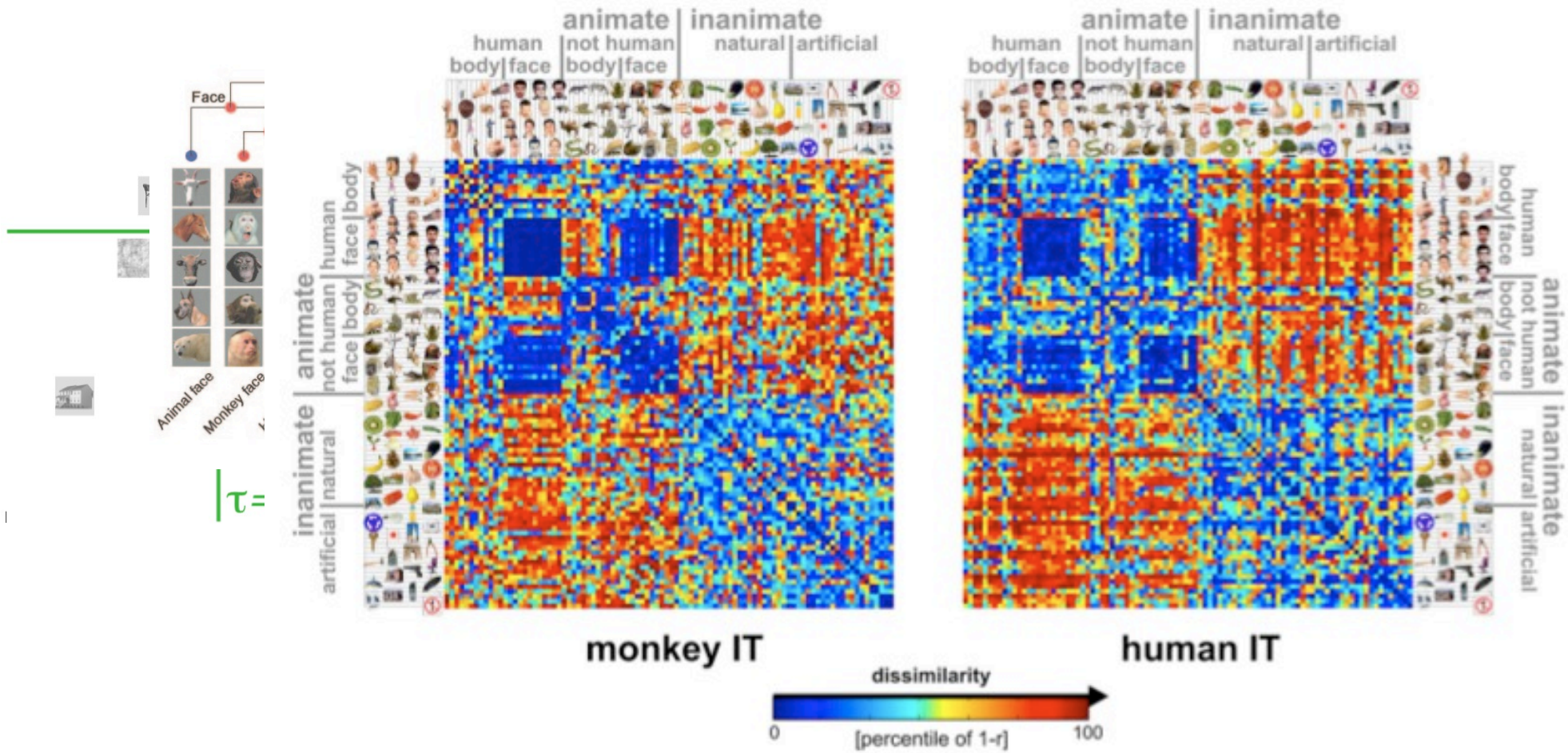
O'Toole et al. 2004

Is the animate-inanimate distinction a major large-scale feature that distinguishes lateral from medial ventral temporal cortex



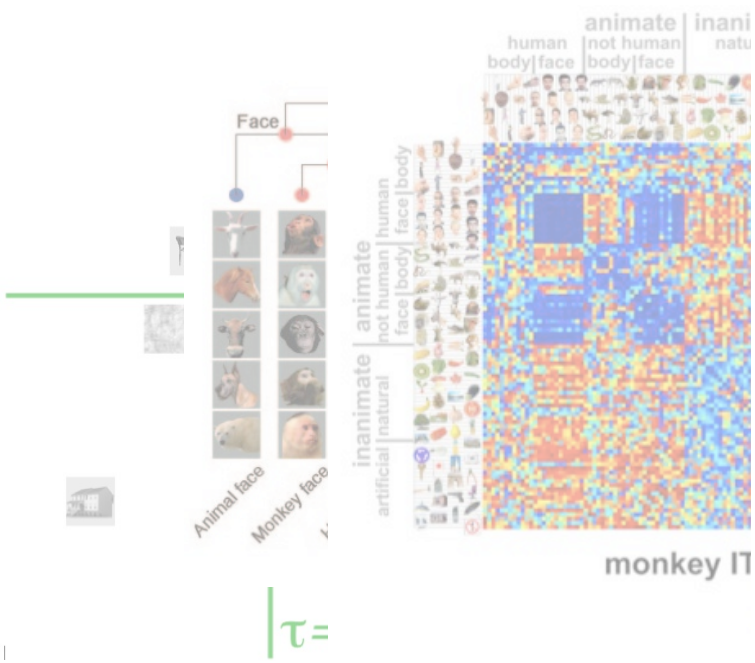
Kiani et al. 2007

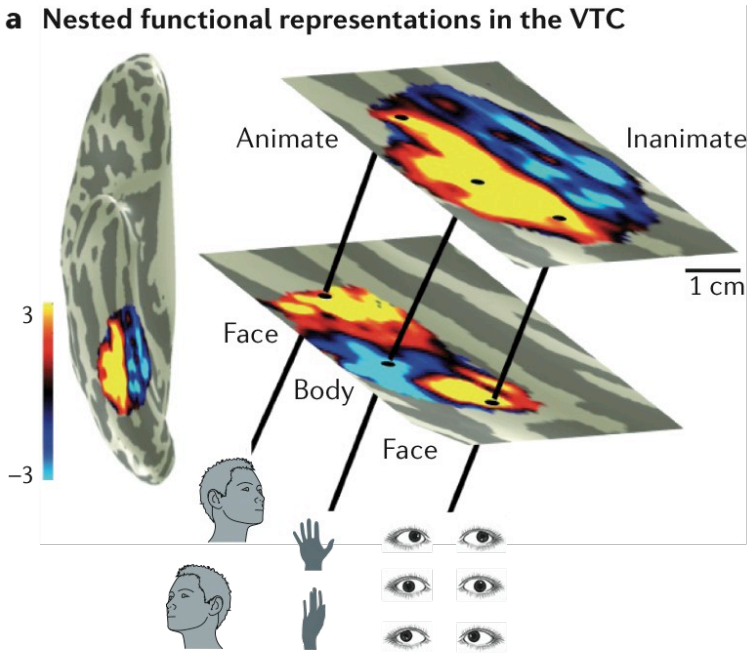
Is the animate-inanimate distinction a major large-scale feature that distinguishes lateral from medial ventral temporal cortex

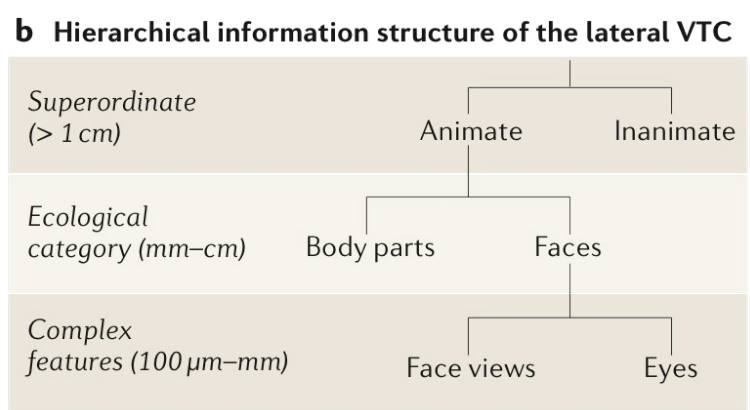


Kriegeskorte et al. 2008

Is the animate-inanimate distinction a major large-scale feature that distinguishes lateral from medial ventral temporal cortex



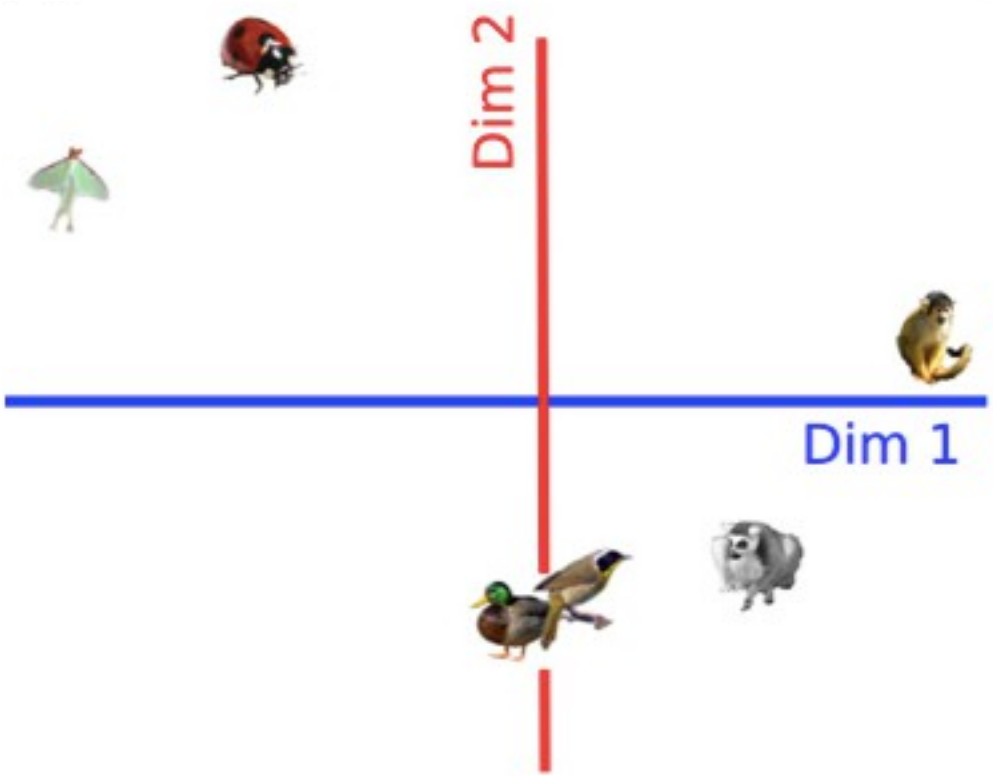




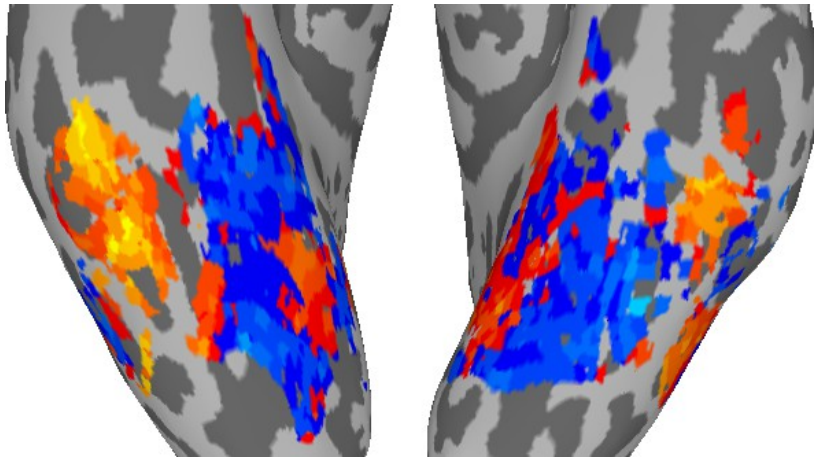
Grill-Spector & Weiner, 2014

But…

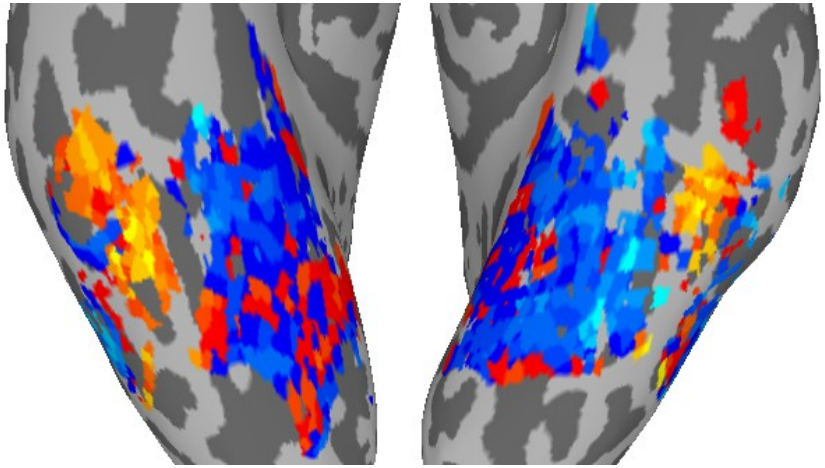
Animate stimuli are usually human, mammal, or avian
Broader sampling suggests this distinction does not apply to all animate stimuli



Connolly et al. 2012

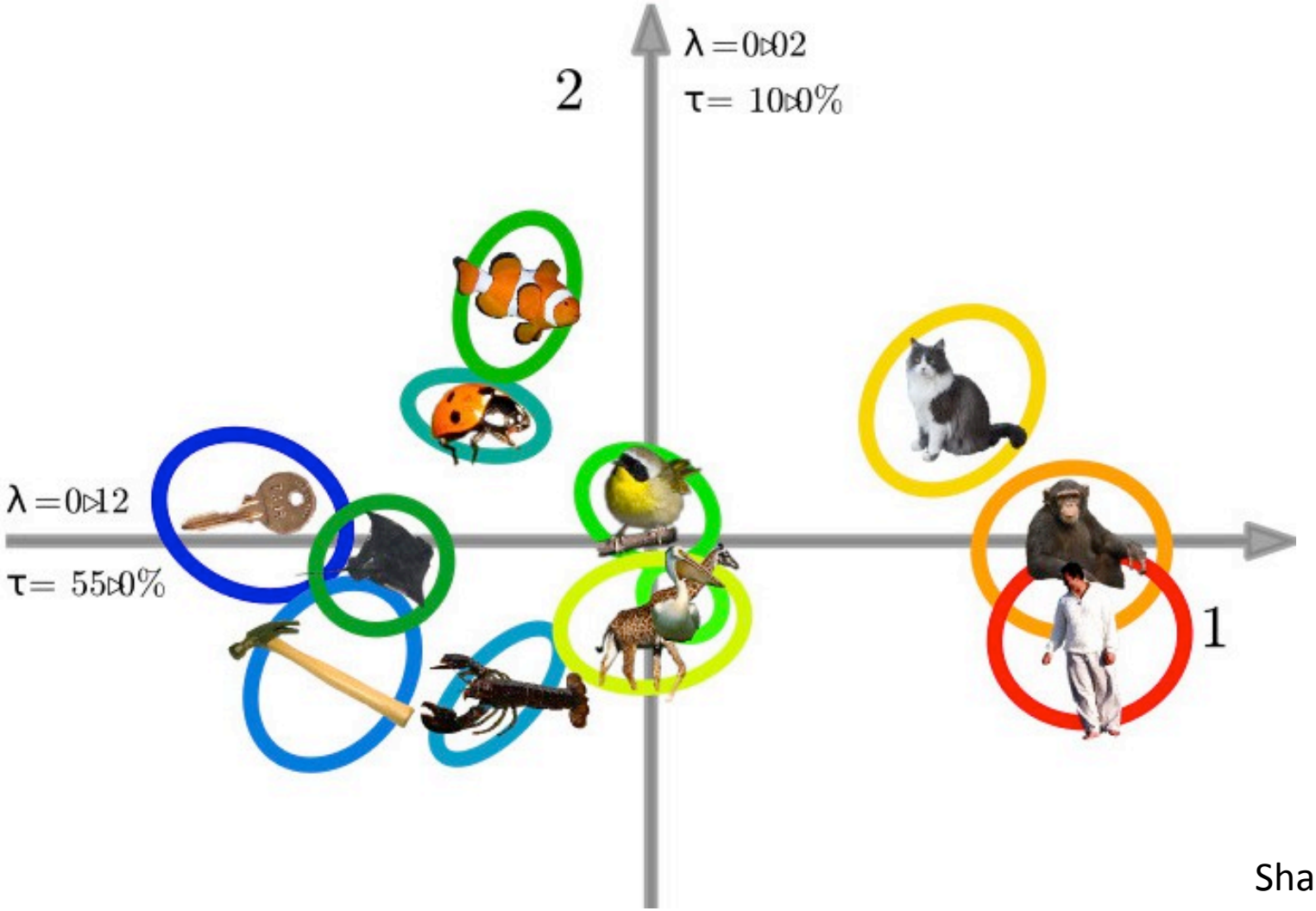


Primates vs bugs



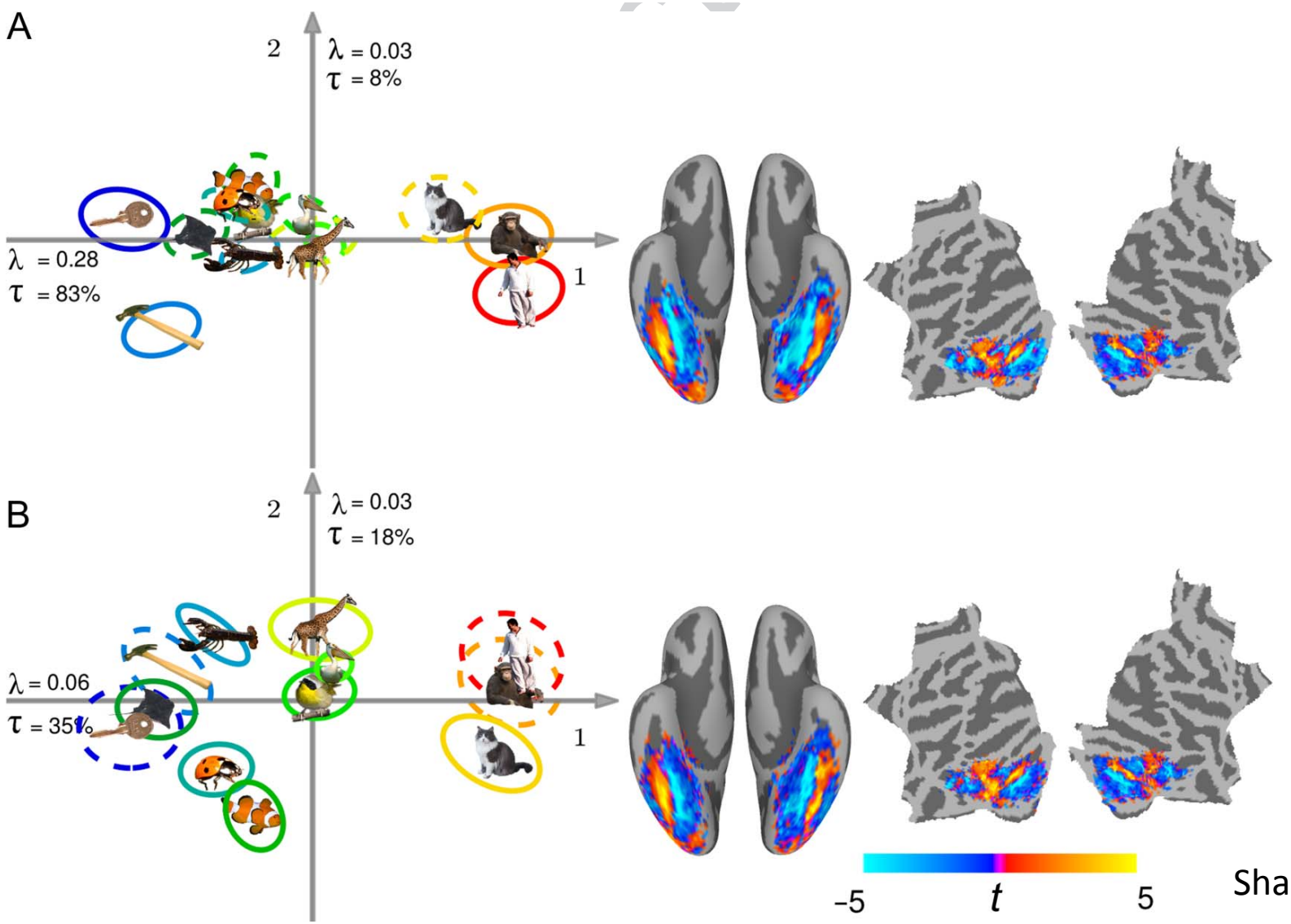
Faces vs objects

The Animacy Continuum:
Low animacy animals have neural response vectors that overlap with vectors for bona fide objects



Sha et al. 2015

The Animacy Continuum:
Topography for high versus low animacy mirrors primates versus objects (animate vs inanimate)



Sha et al. 2015

Animate-Inanimate dichotomy does not survive broadening stimulus sampling to ‘low animacy’ animals

The animate-inanimate dichotomy is a special case of a broader principle – the *Animacy Continuum*.

c.f. Aristotle, *De Anima*; animacy hierarchies in linguistics; the “Great Chain of Being”

Animate-Inanimate dichotomy does not survive broadening stimulus sampling to 'low animacy' animals

The animate-inanimate dichotomy is a special case of a broader principle – the *Animacy Continuum*.

c.f. Aristotle, *De Anima*; animacy hierarchies in linguistics; the "Great Chain of Being"

So… Is the *Animacy Continuum (varying levels of agentic complexity)* the dominant organizing principle in VT cortex?

Representation of agents

Gobbini et al. 2007

"…these results suggest that the fusiform gyrus plays a general role in the representation of visual stimuli that signify agency, independent of visual form."

Gobbini et al. 2011

"These results suggest that these areas mediate perception of agency, independently of whether the agents are living or not."

Shultz and McCarthy, 2014

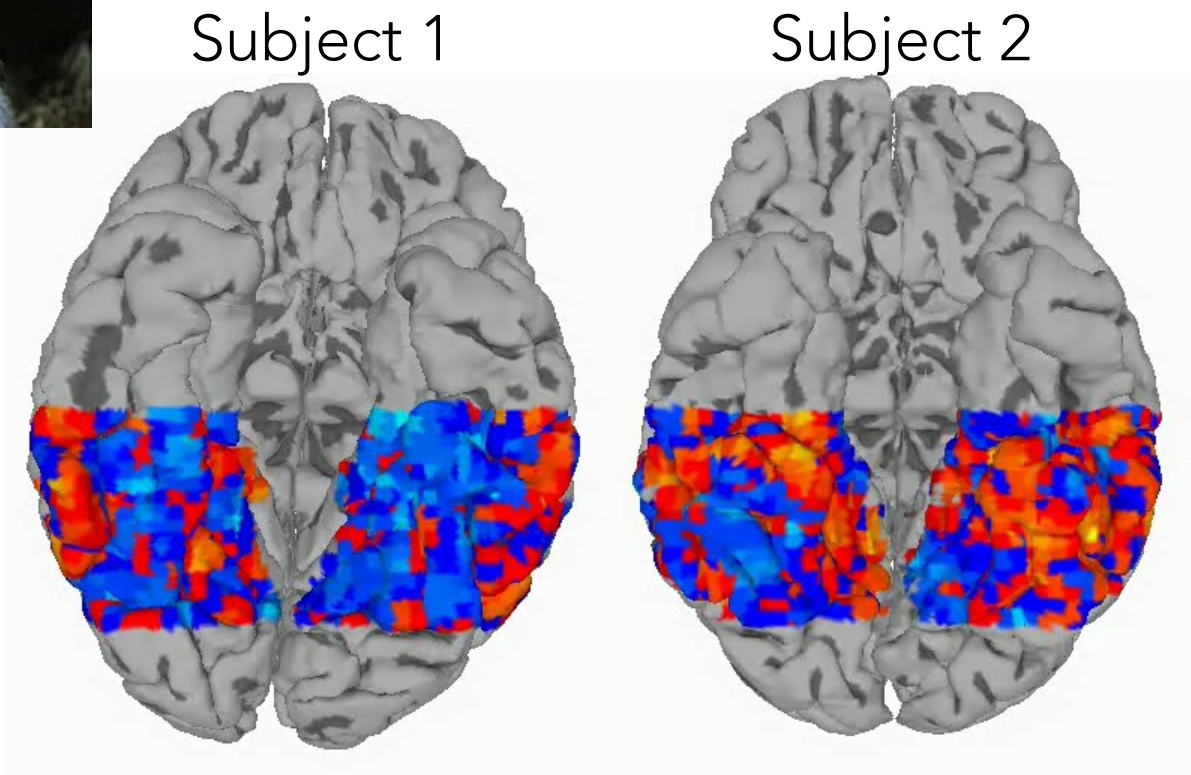
While decades of research have demonstrated that a region of the right fusiform gyrus (FG) responds selectively to faces, a second line of research suggests that the FG responds to a range of animacy cues, including biological motion and goal-directed actions, even in the absence of faces or other human-like surface features. These findings raise the question of whether the FG is indeed sensitive to faces or to the more abstract category of animate agents… [Our] results suggest that the FG does not respond to all faces in a category-specific way, and is instead especially sensitive to whether an entity is animate.

… Is the animacy continuum (varying levels of agentic complexity) *the dominant organizing principle* for representation in ventral temporal cortex?

Not exactly

Time for a digression to reframe the question

Methods for modeling a high-dimensional representational space based on a set of basis functions (tuning profiles, connectivity profiles, and topographic components) that are shared across brains – Hyperalignment

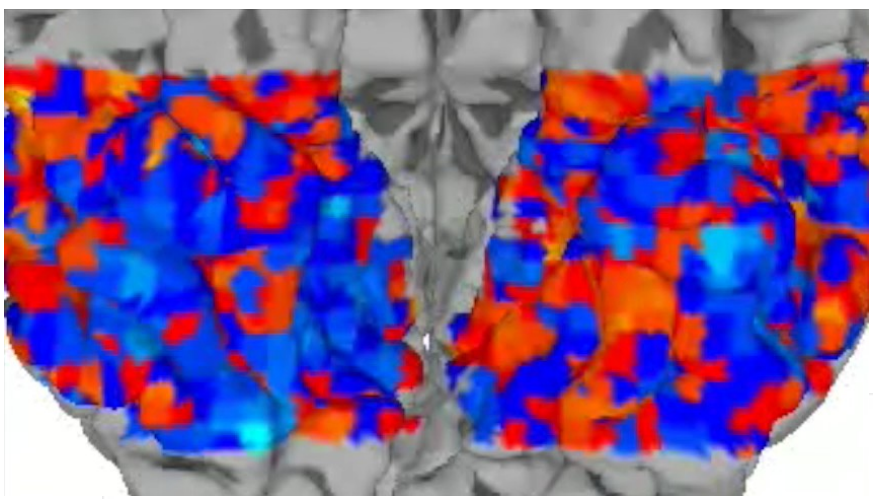

Subject 1
Subject 2

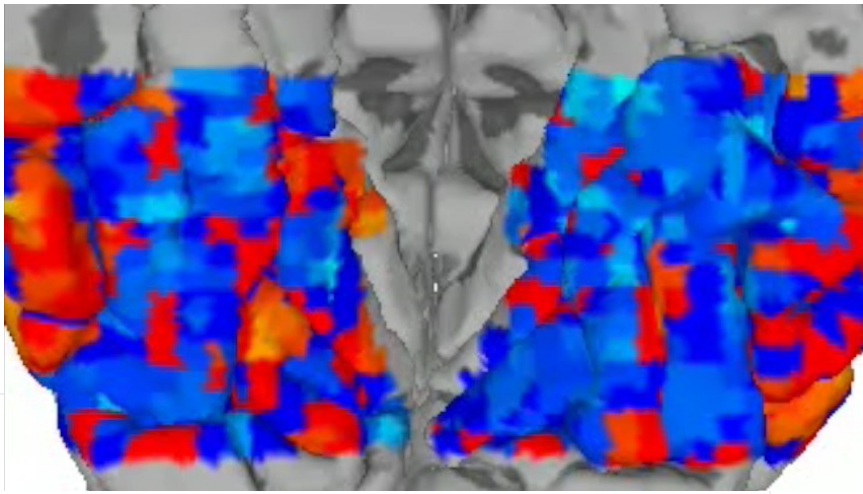
Broad sampling of a neural representational space with a movie

Subject 1

Subject 2

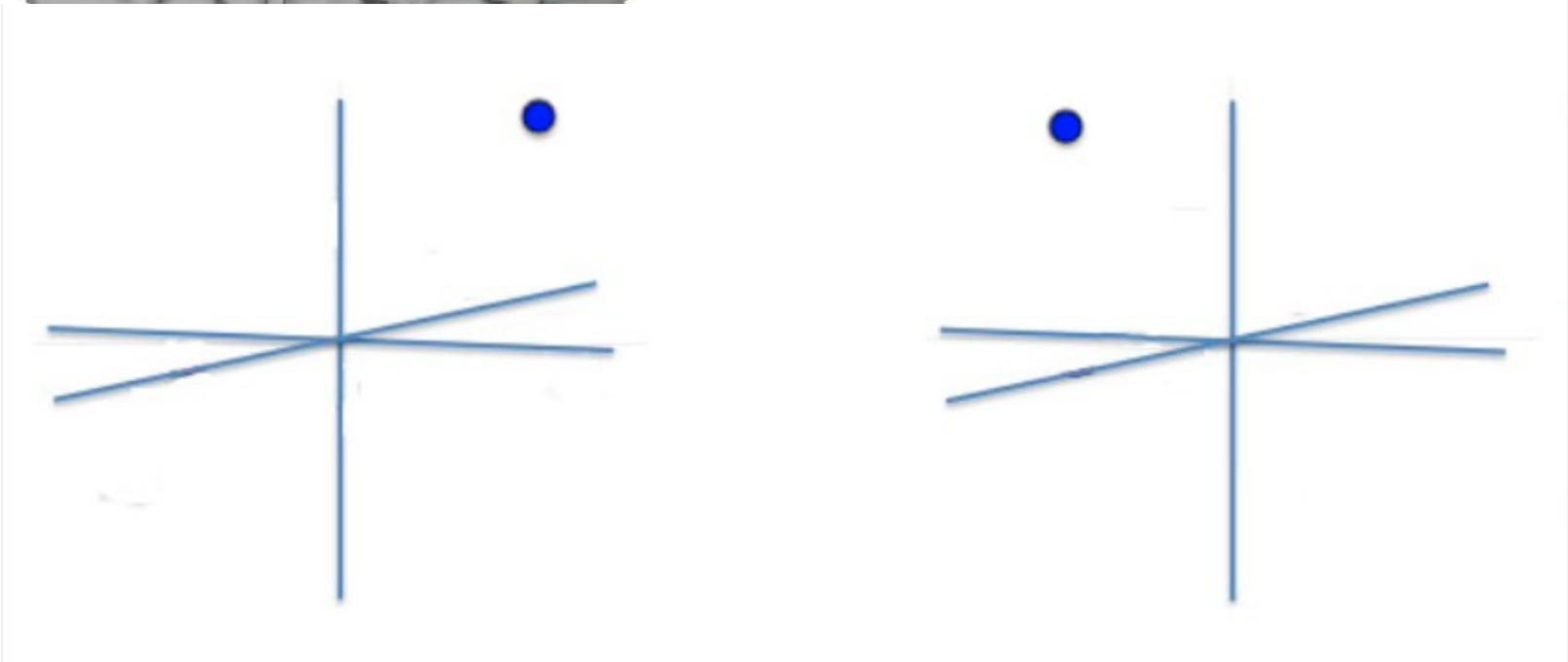
Response patterns in cortex

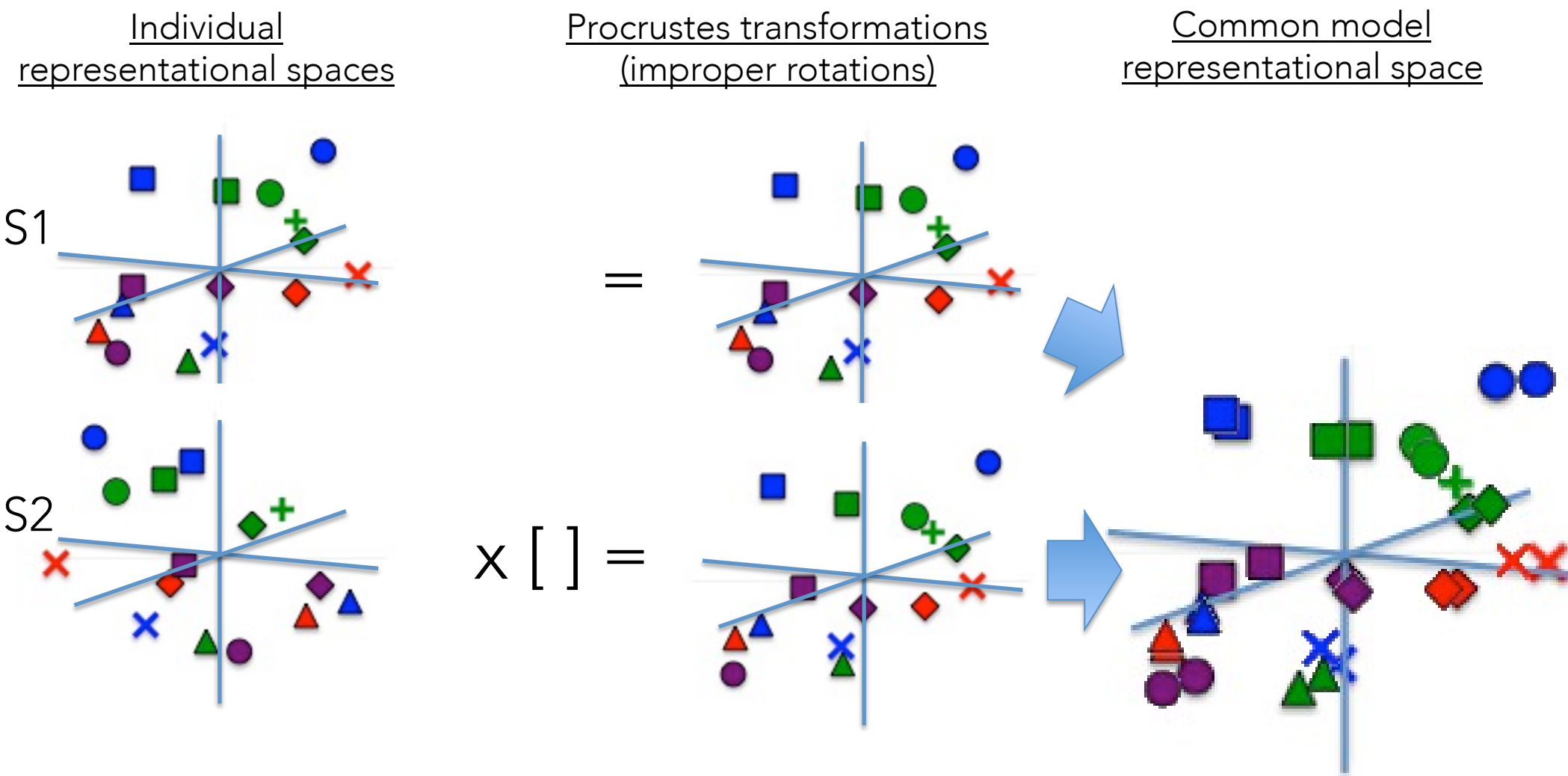


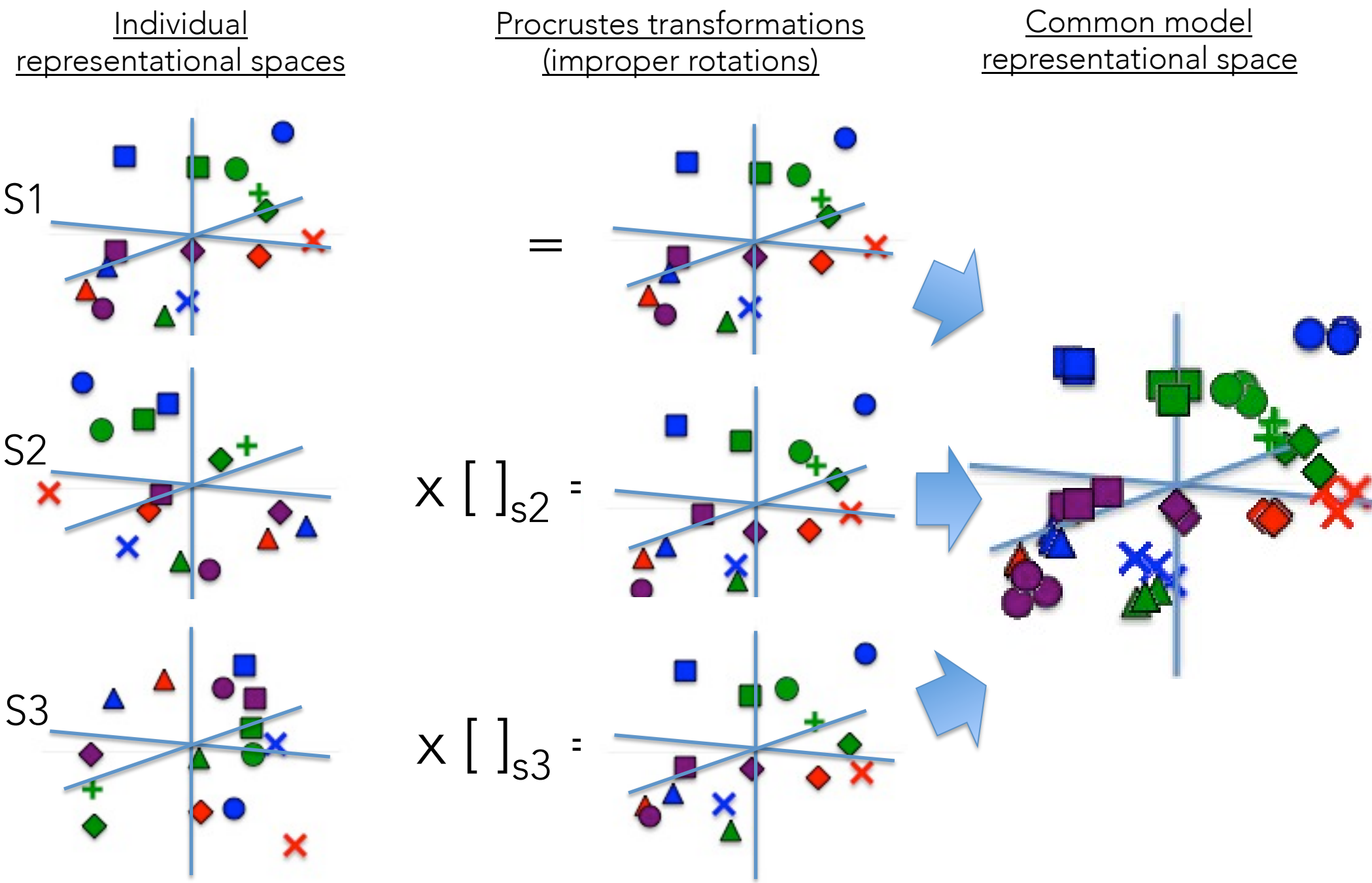


15 response pattern vectors in individual 3D representational spaces

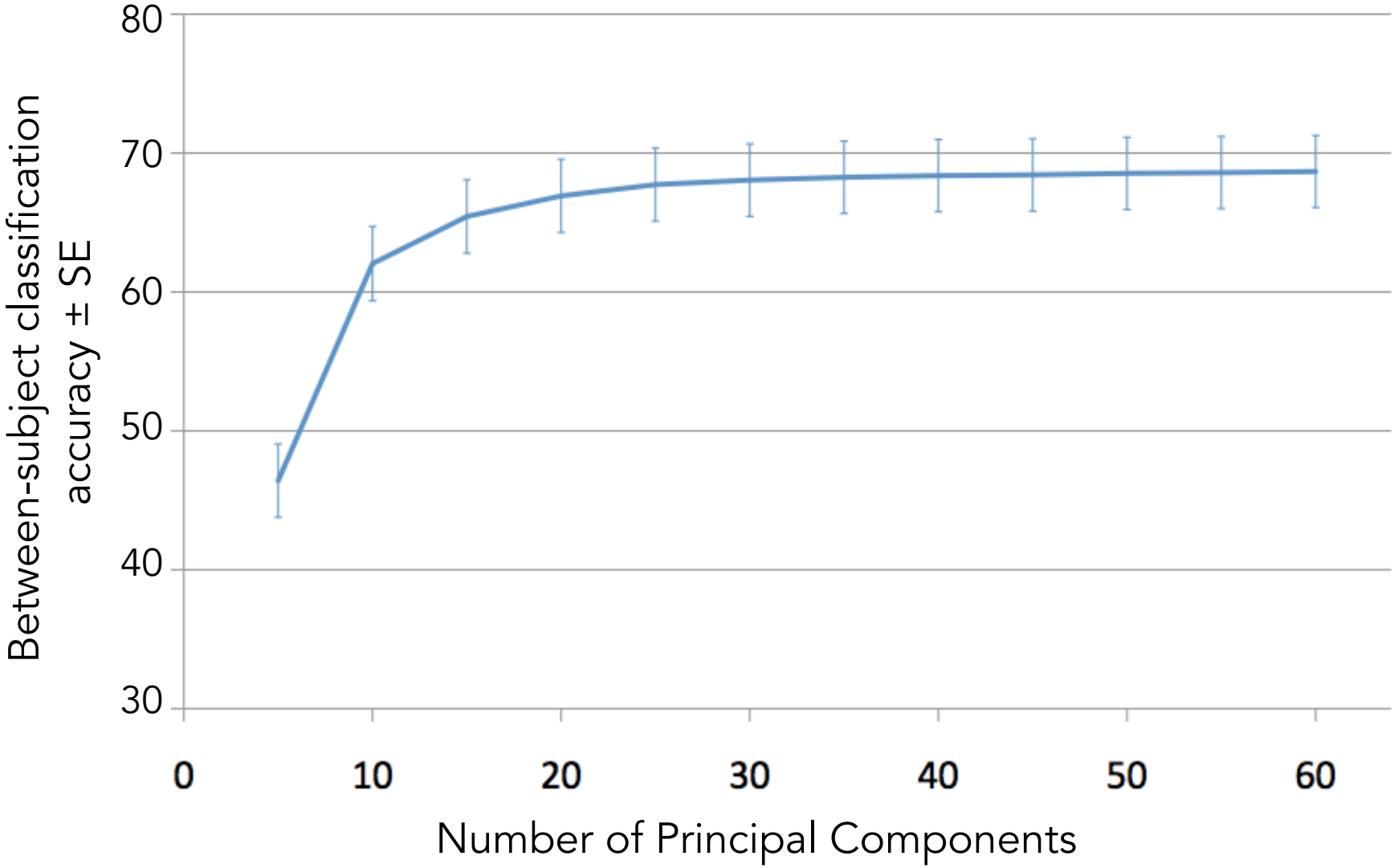
(full exp't has >2600 vectors in >50,000D space)



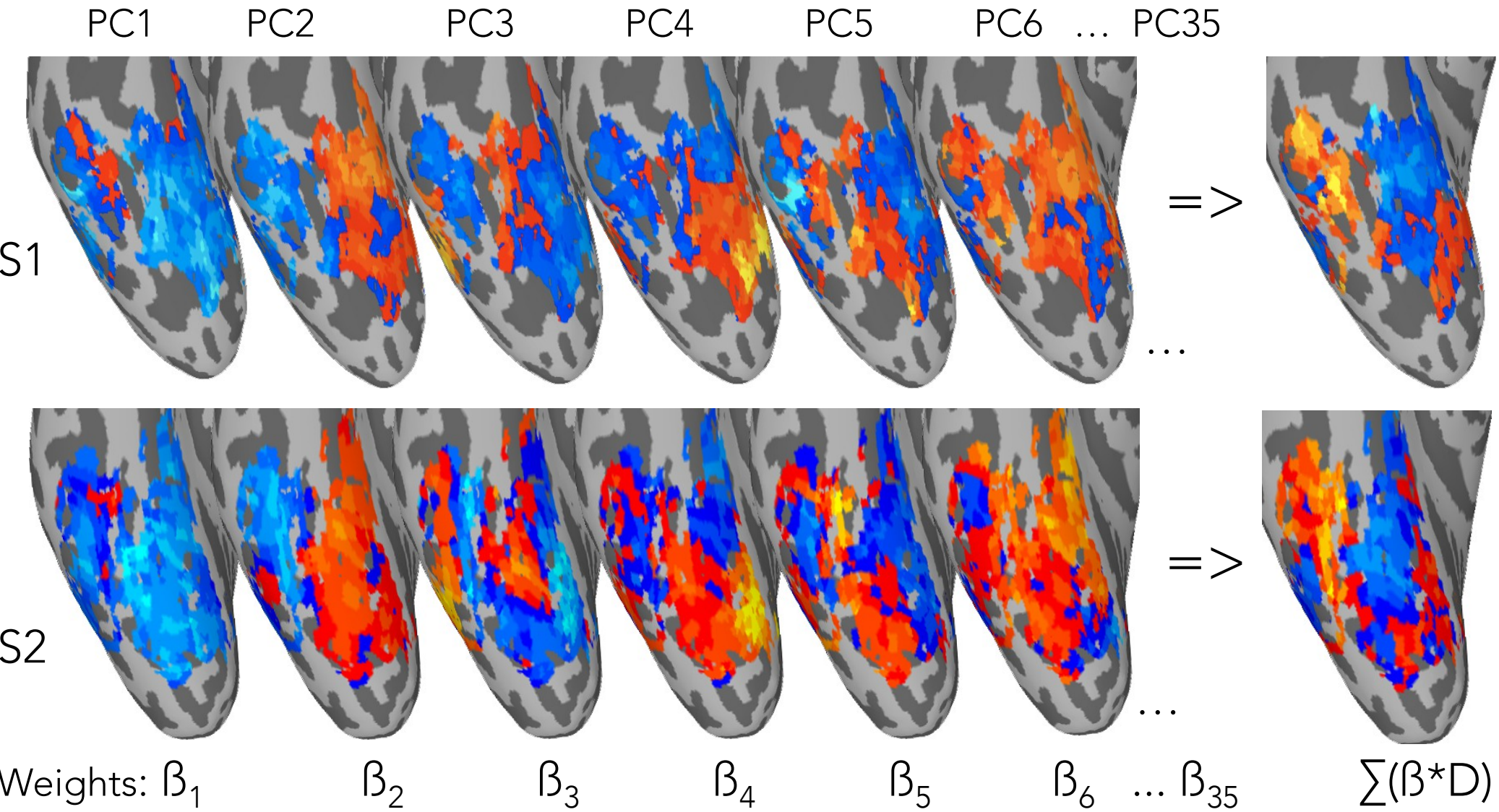

Individual representational spaces
Procrustes transformations (improper rotations)
Common model representational space
S1
S2
=
x [ ] =


Individual representational spaces
Procrustes transformations (improper rotations)
Common model representational space
S1
S2
S3
=
x [ ]s2 =
x [ ]s3 =

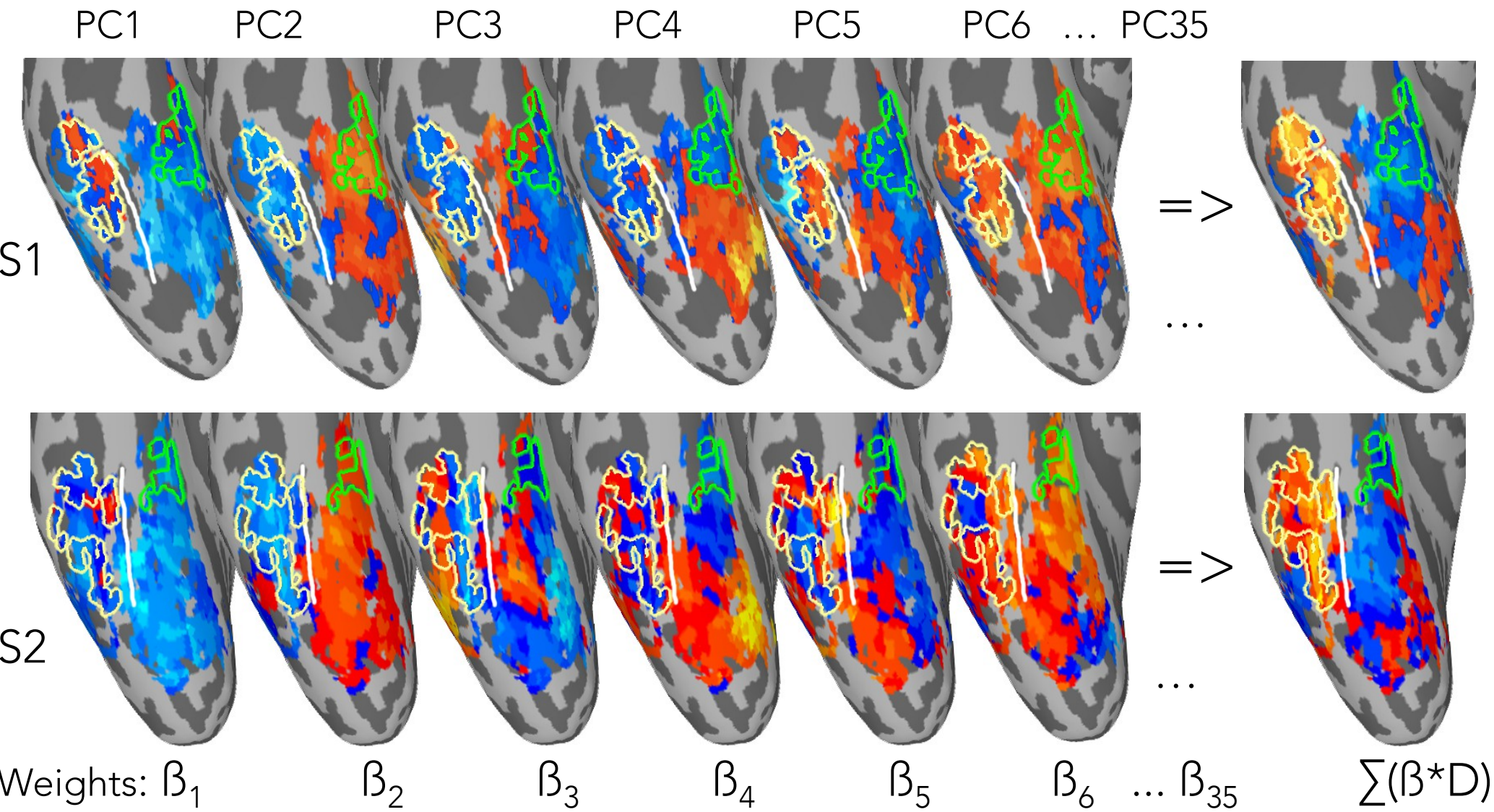
>30 dimensions (42?) are necessary to account for the representational space in VT



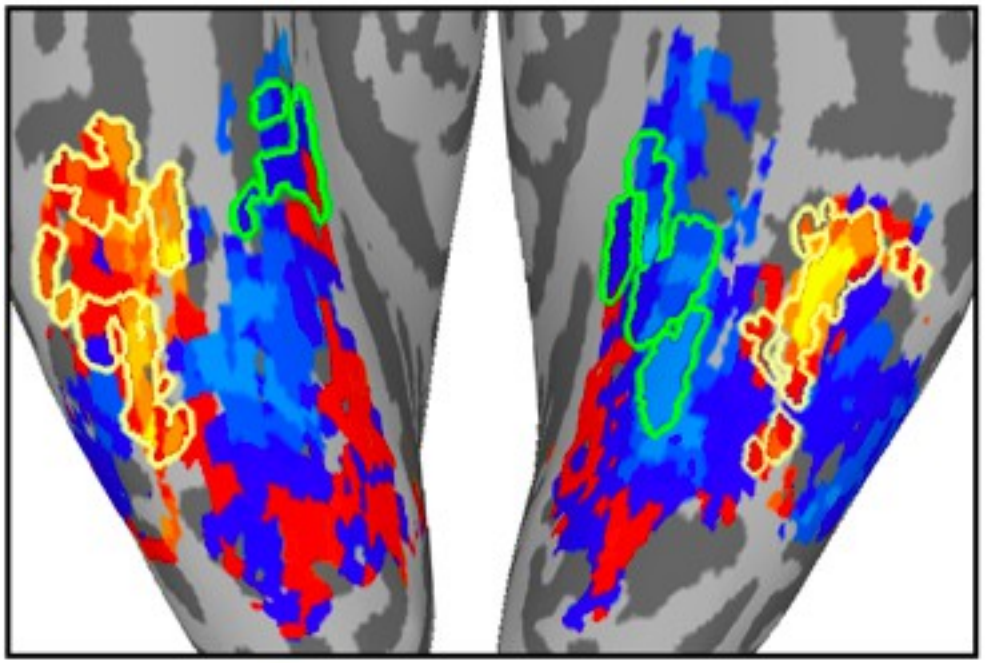
Topographies for response patterns are modeled in different brains as weighted sums of individual-specific topographic basis functions using <u>the same weights for common model dimensions</u>



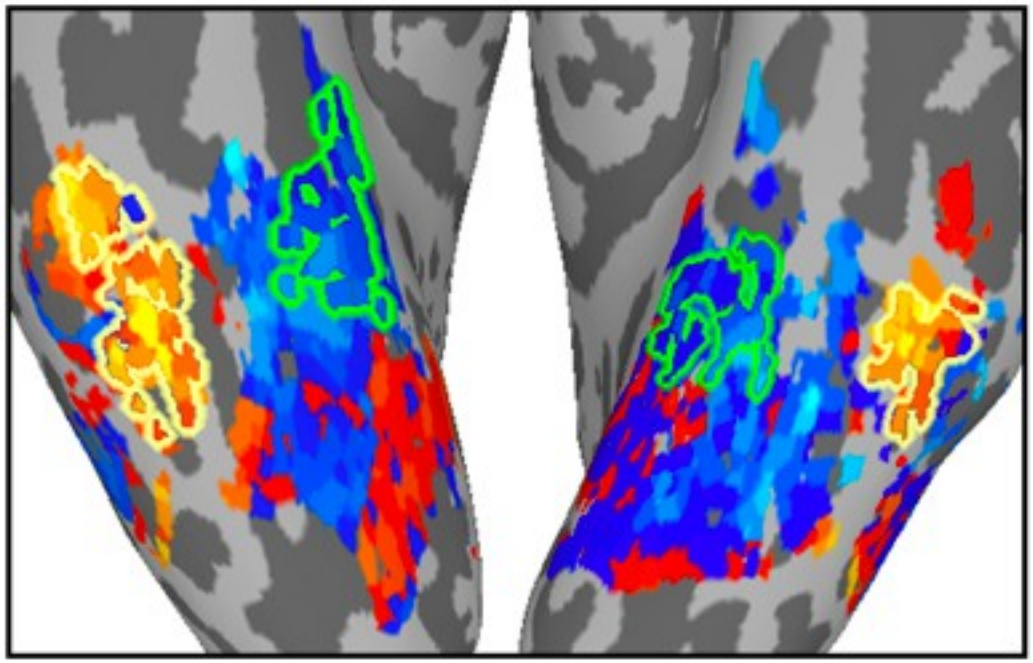
The topographic basis functions for PCs individually show little correspondence to category-selective face and place areas or the domain-specific divisions for animate and inanimate stimuli



Individual variability in the location of the FFA is modeled by VT model topographies for face vs object dimension with high fidelity

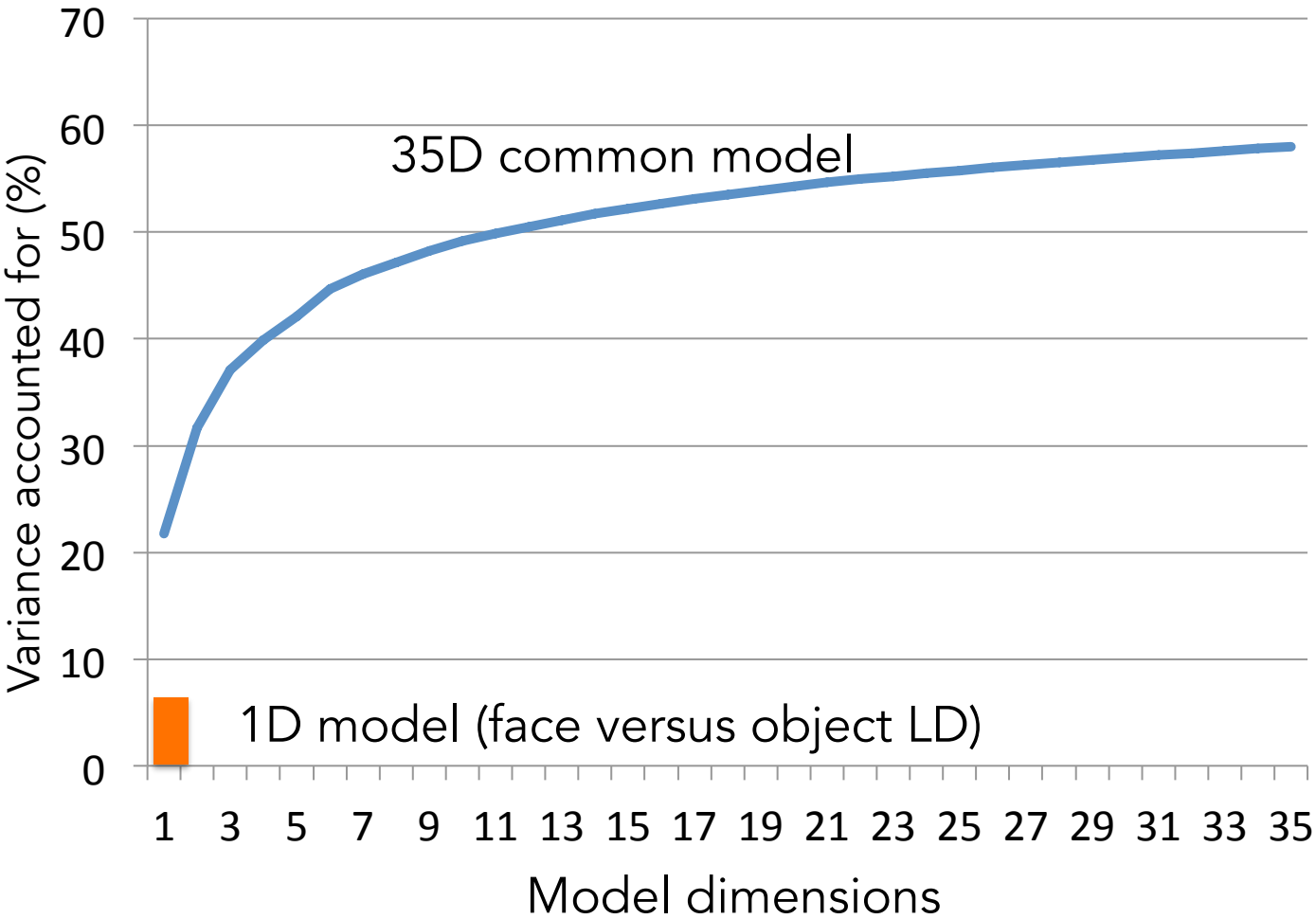


Subject 1



Subject 2

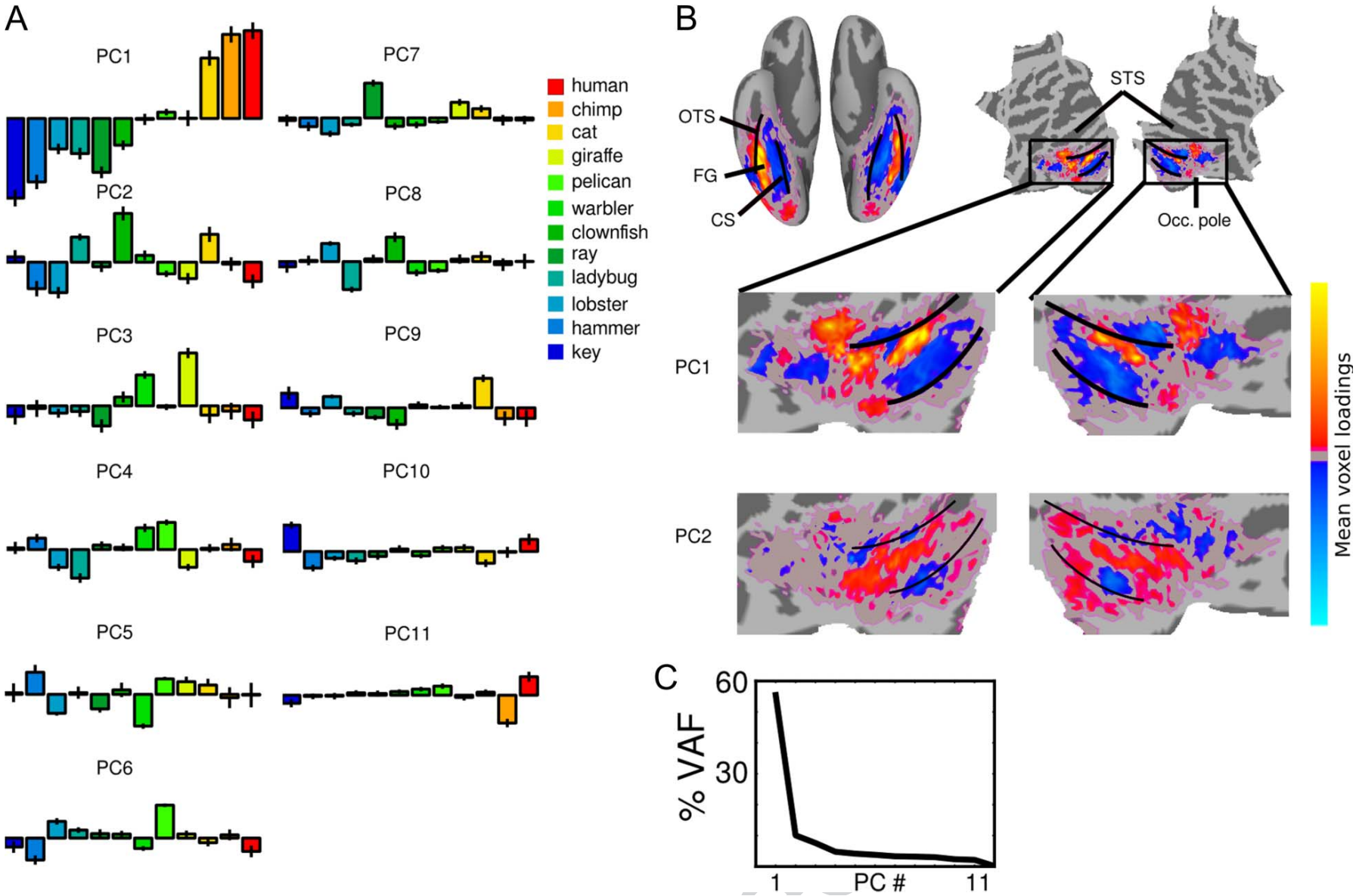
But this dimension accounts for only a small portion of variance in *responses to the movie*



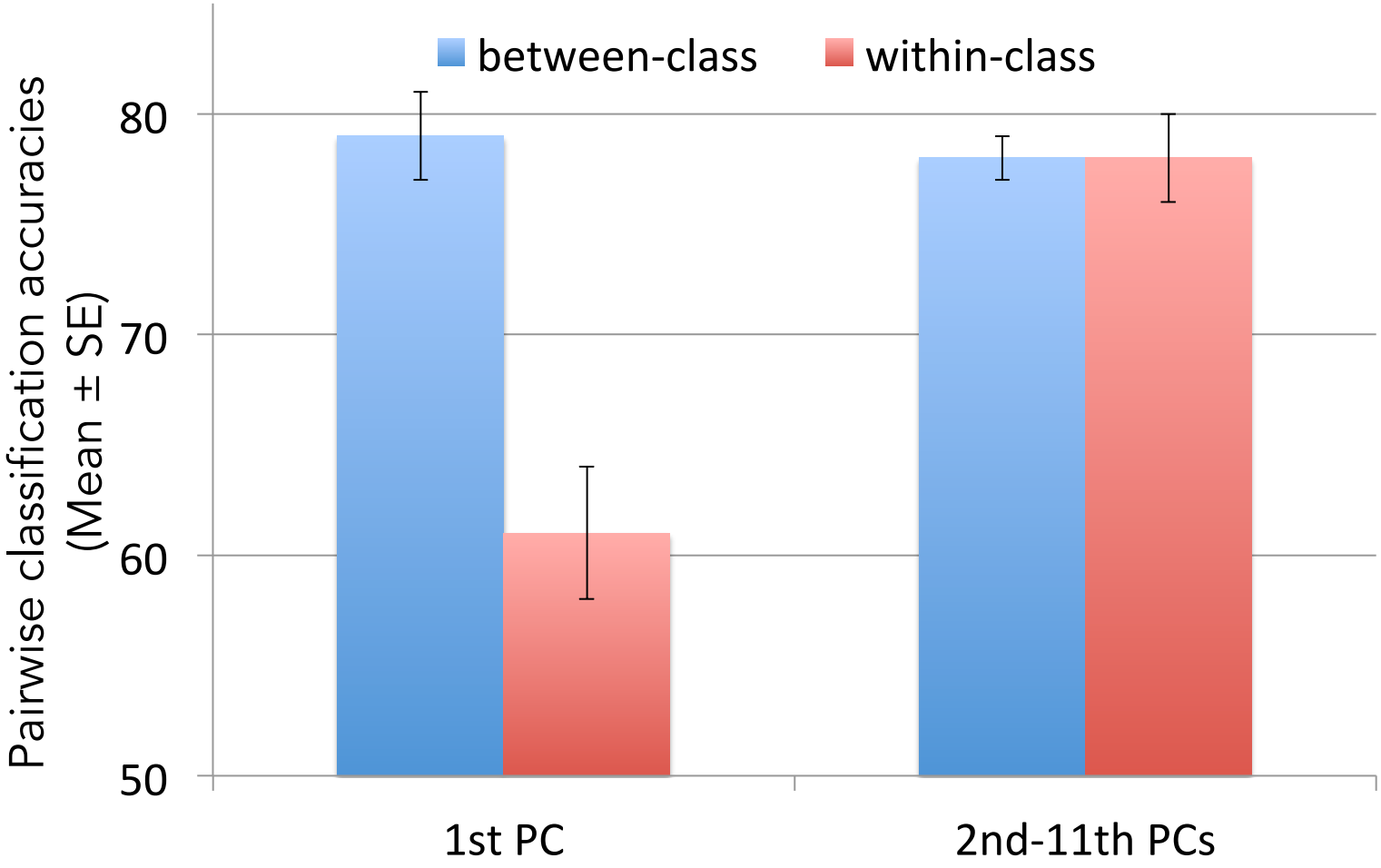
Face vs object LD accounts for 7% of variance, <20% of VAF by first 3 PCs & <13% of VAF by 35 PCs

The animacy-continuum/face-object dimension accounts for more variance in responses to still images of animals and objects

(Sha et al. 2015)



…but only carries information about coarse distinctions – lower order dimensions carry finer distinctions



Sha et al. 2015

… Is the animacy continuum (varying levels of agentic complexity) *the dominant organizing principle* for representation in ventral temporal cortex?

Not exactly

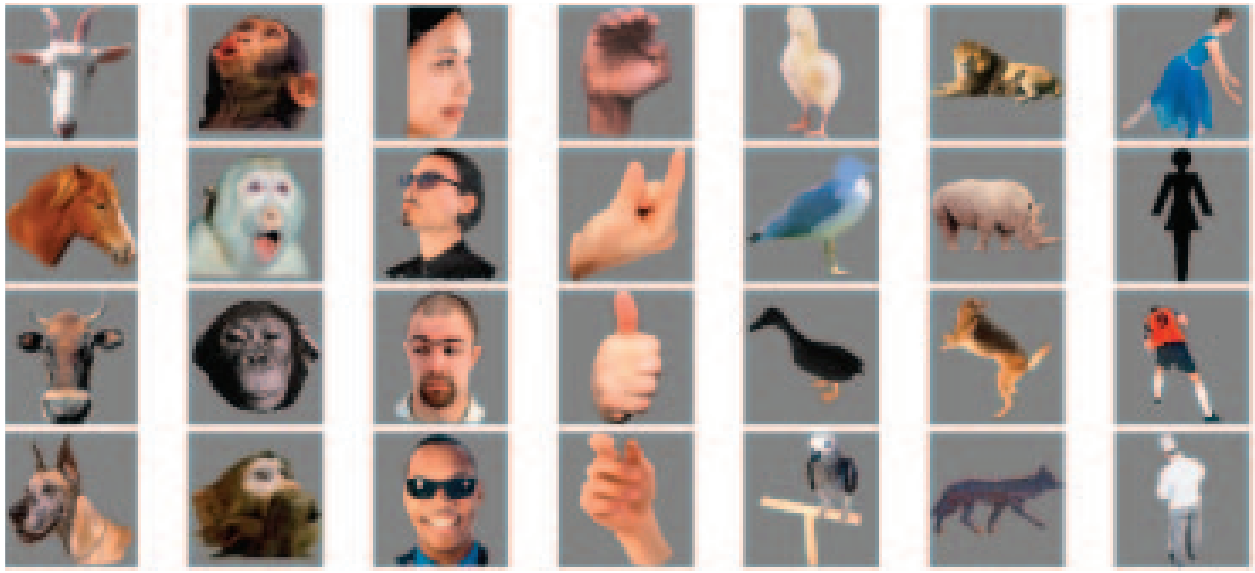
The animacy continuum and the face-object distinctions account for only a small portion of variance in the responses to a rich, natural, dynamic stimulus

The animacy continuum and face-object distinctions account only for coarse, not fine-grained distinctions

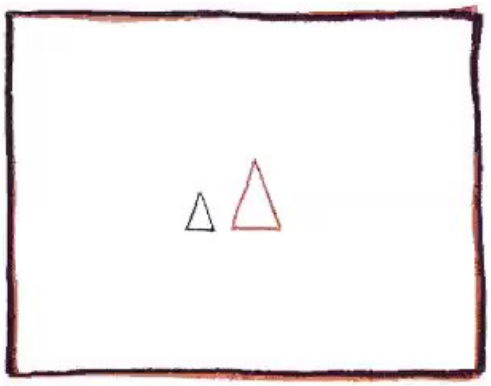
But the dimensions for these distinctions are derived from responses to still images. Dimensions based on responses to agentic behavior may tell a different story (cf. Nastase et al.)

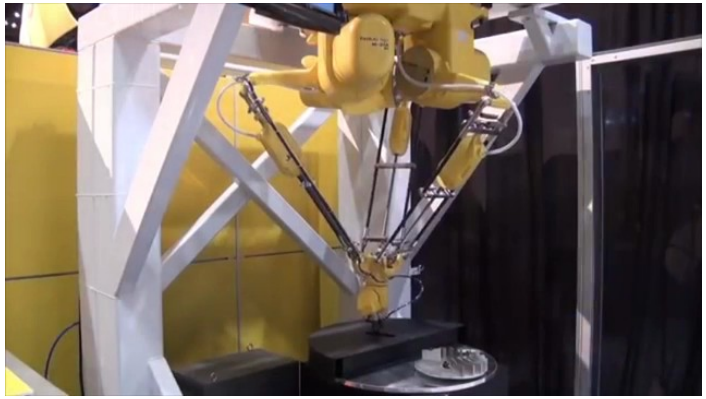
Summary 1: Stimulus sampling bias

Oversampling of still images from limited variety of categories (faces, bodies, mammals, birds, etc.)



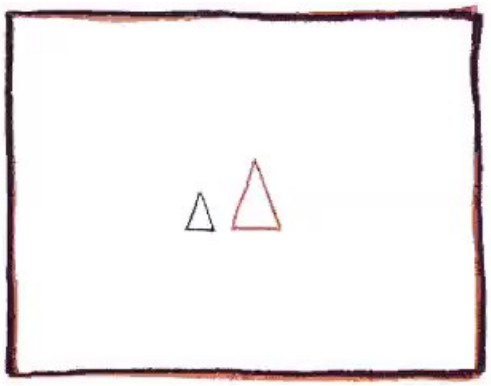
versus

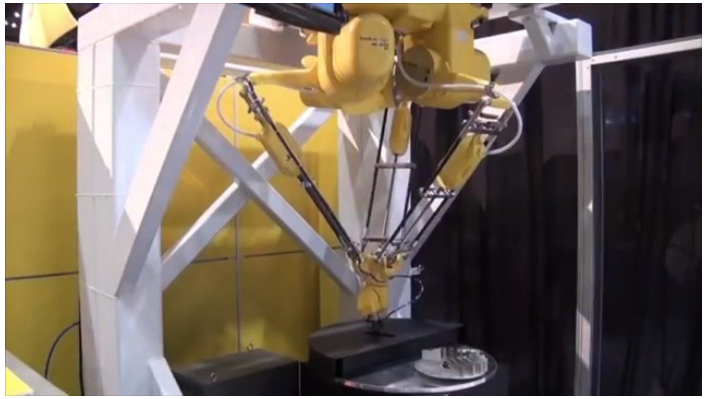




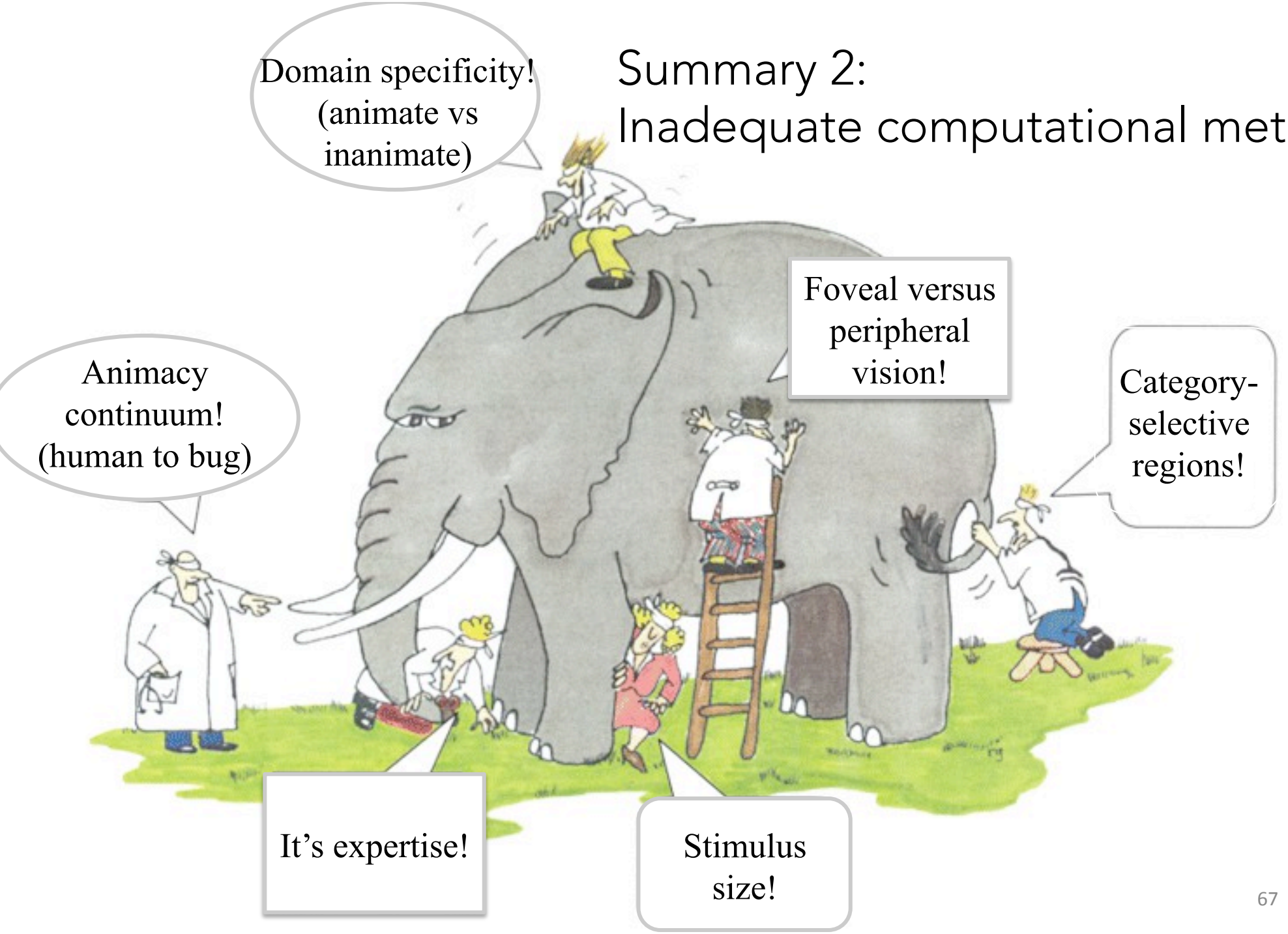
Summary 1

- "Category selective" regions have more complex tuning functions that suggest they play more diverse roles in person perception
- Representational geometry is dominated by representation of behavior, not form
- The animate-inanimate distinction is not the principal dimension that characterizes the coarse lateral-to-medial topography in ventral temporal cortex





Summary 2:
Inadequate computational methods



Summary 2:
Inadequate computational methods

Univariate statistics

- Limiting analysis to univariate contrasts leads to search for the 'function' of an area

Multivariate methods (pattern classification, RSA, hyperalignment, etc.)

- Model the functional architecture of an area as a *high-dimensional representational space*
- The topographies for dimensions are multiplexed and overlapping
- Accounts for category-selective regions and fine-grained patterns that carry fine distinctions

Thank you

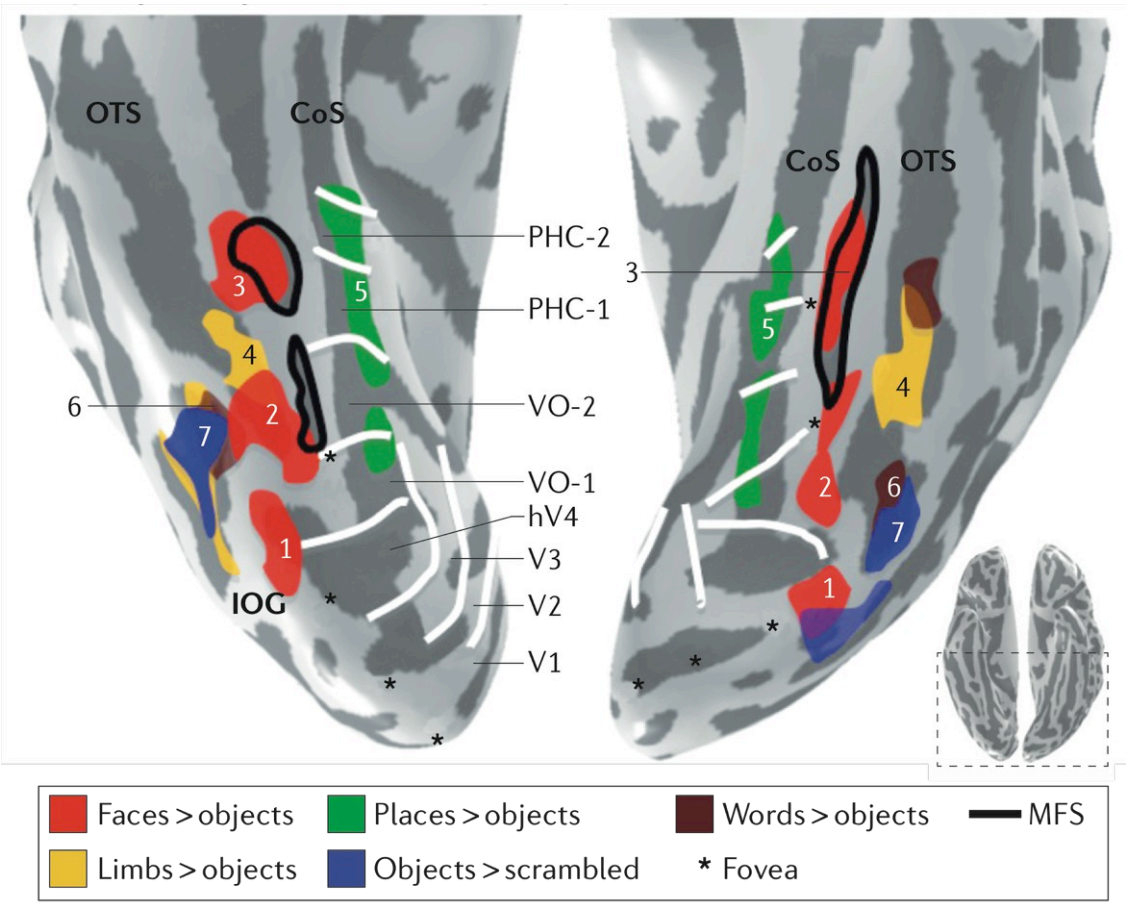
Software for ROI hyperalignment and data are on PyMVPA (www.pymvpa.org)

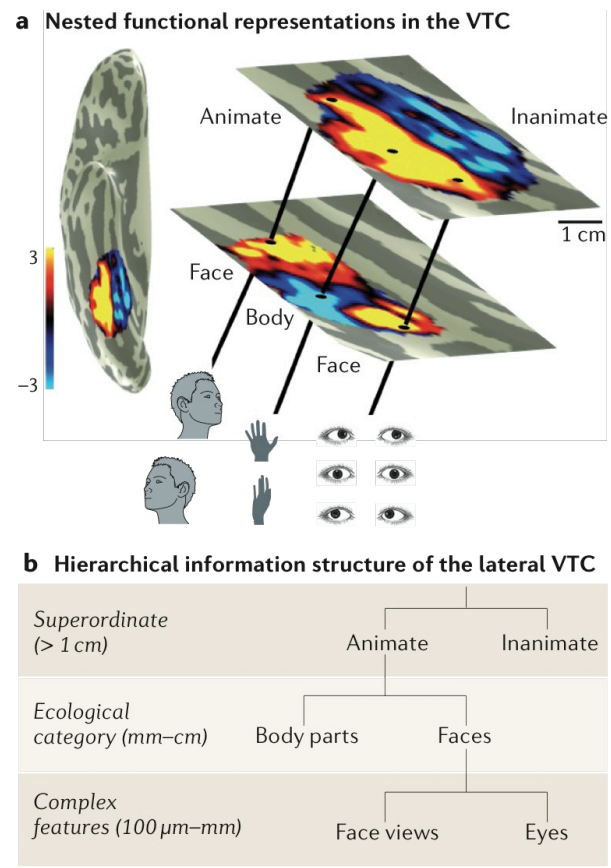
Yaroslav Halchenko & Michael Hanke, developers

New massive data release of 7T fMRI with natural stimulus and lots more:

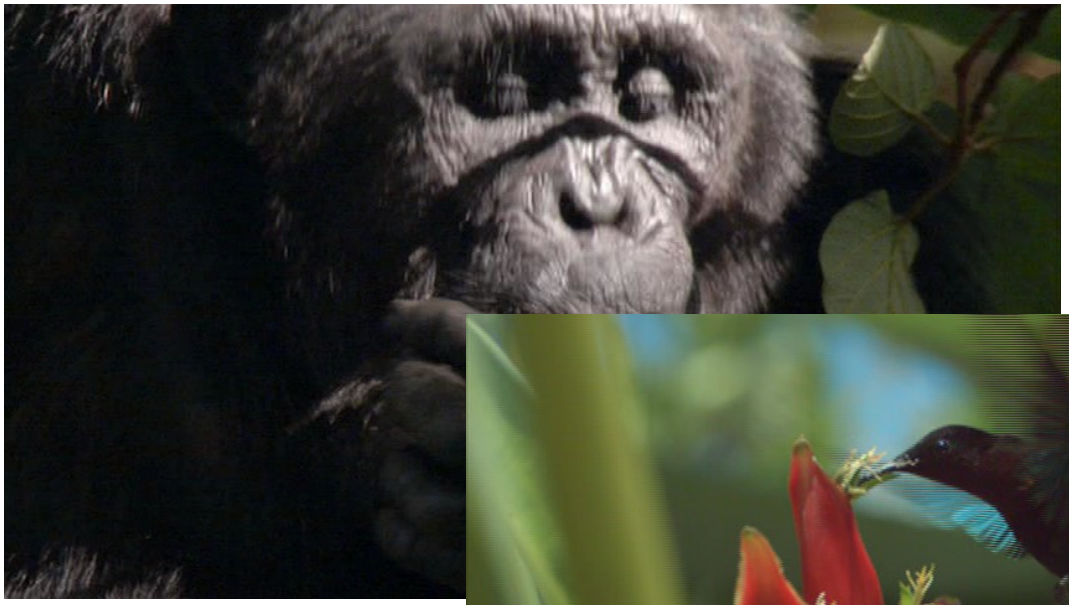
data website: http://www.studyforrest.org

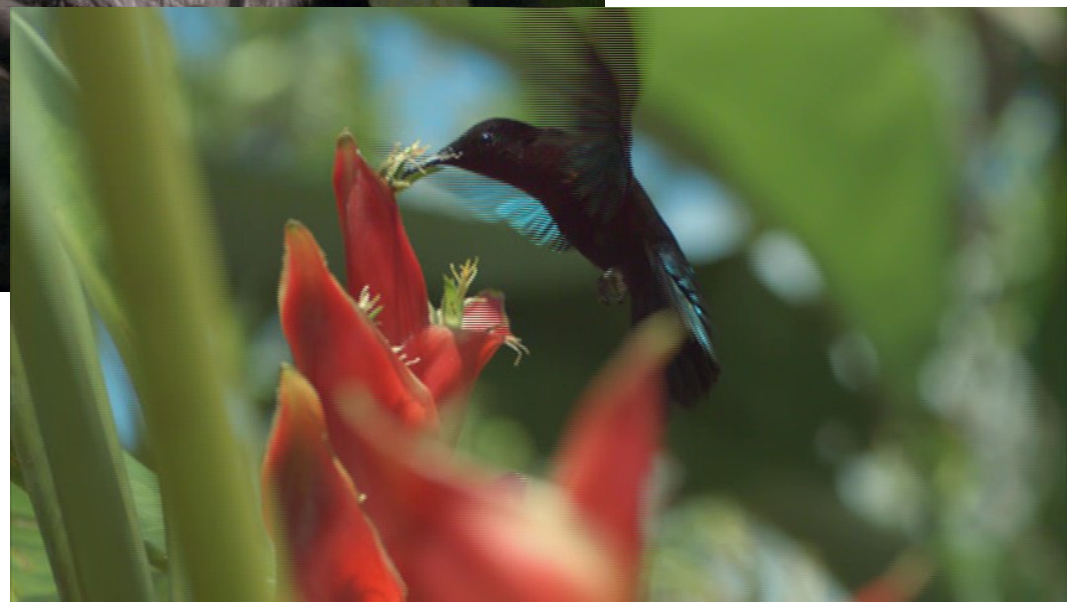
paper: Hanke et al. (2014) *Nature Scientific Data, 1: 140003.*





Grill-Spector & Weiner, *Nature Reviews Neuroscience*, 2014





Thank you