Differential neural responses in high-level visual cortex explain categorization behavior

Mona Rosenke
N. Davidenko, K. Grill-Spector, K. S. Weiner

Vision and Perception Neuroscience Lab
Stanford University

How do we categorize ambiguous images?

How do we categorize ambiguous images?

What is this?

A Caribbean island?

A face?

A hand?

How does visual cortex support this categorization?

What is this?

A Caribbean island?

A face?

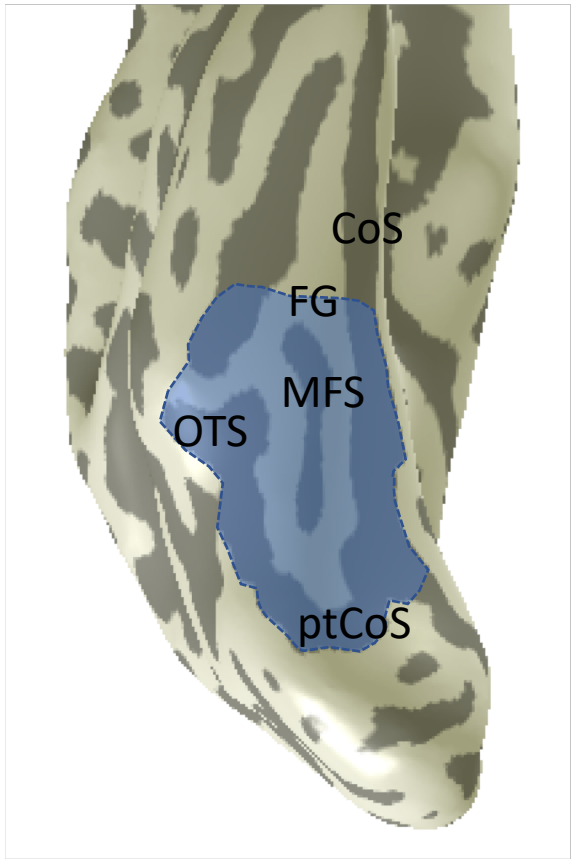
A hand?

A long line of research shows that ventral temporal cortex supports the perceptual ability to categorize visual images

Brain directions: Inflated cortex with view of the ventral occipital and temporal lobe

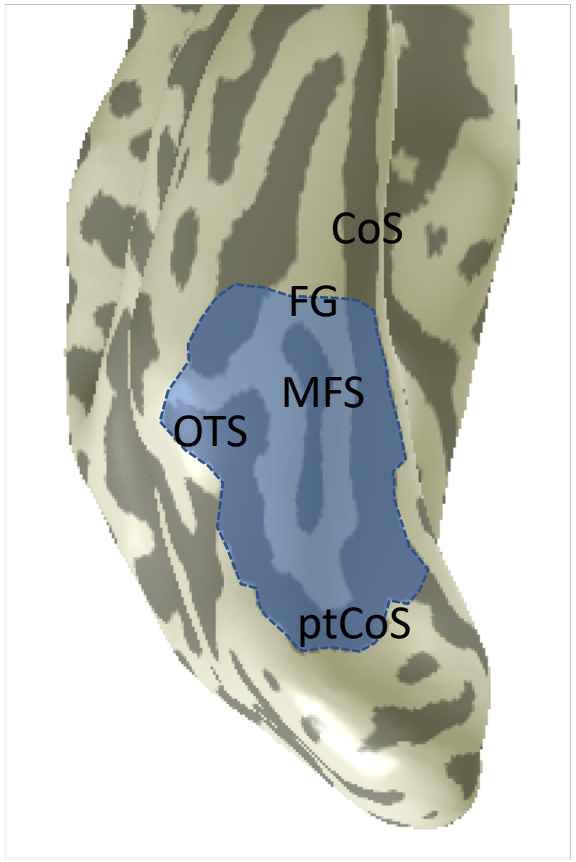
dark gray = sulci
light gray = gyri

Ventral temporal cortex (VTC) is an anatomical area that can be defined by a series of anatomical landmarks



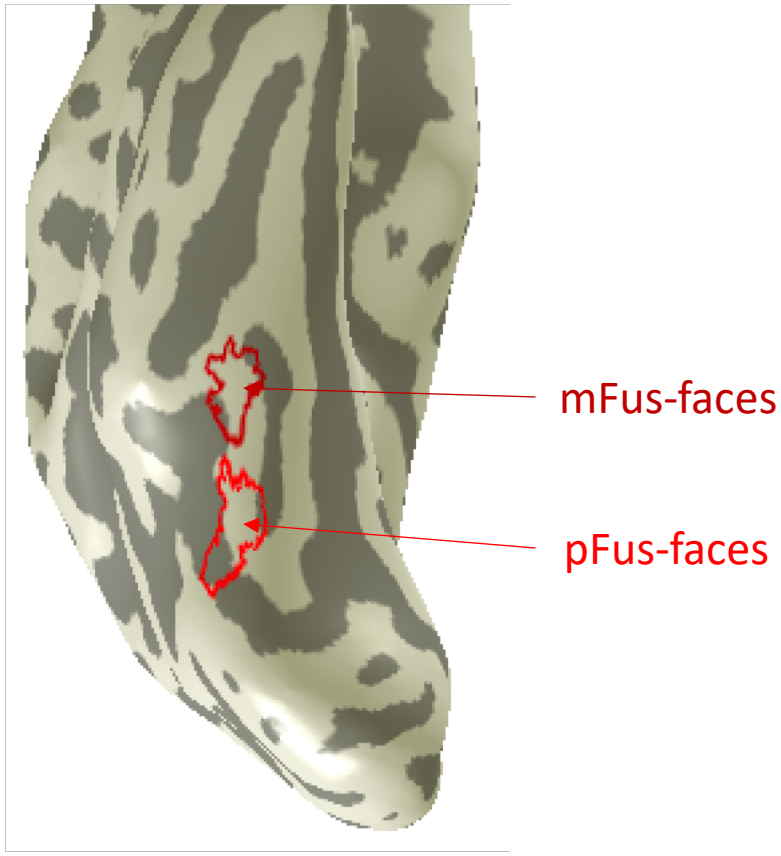
CoS = Collateral sulcus
FG = Fusiform gyrus
MFS = Mid-fusiform sulcus
OTS = Occipital temporal sulcus
ptCoS = posterior transverse collateral sulcus

VTC achieves efficient categorization by having small specialized clusters for different visual categories – e.g. faces



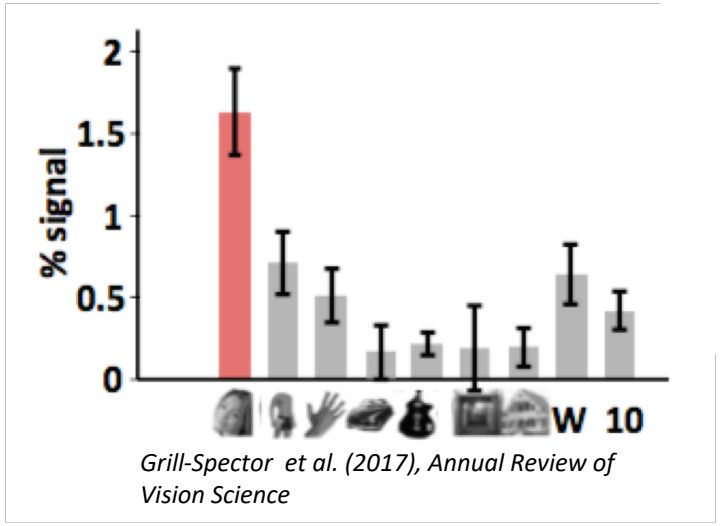
CoS = Collateral sulcus
FG = Fusiform gyrus
MFS = Mid-fusiform sulcus
OTS = Occipital temporal sulcus
ptCoS = posterior transverse collateral sulcus

VTC achieves efficient categorization by having small specialized clusters for different visual categories – e.g. faces



faces > other categories, t = 3

These regions are referred to as face-selective because they exhibit higher fMRI signals to faces than other stimuli

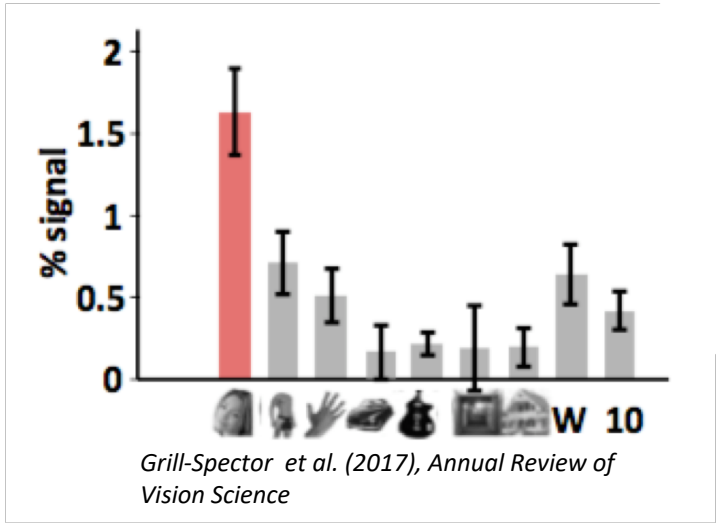


Grill-Spector et al. (2017), Annual Review of Vision Science

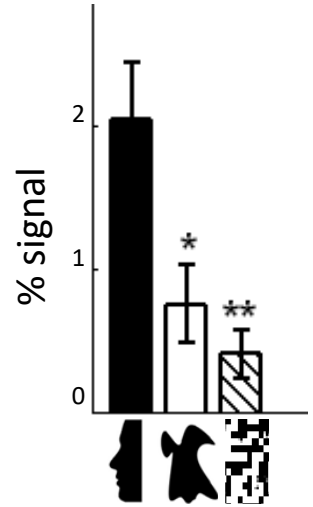
faces > other categories, t = 3

These regions show higher signals for faces across stimulus formats

faces > other categories, t > 3
limbs > other categories, t > 3

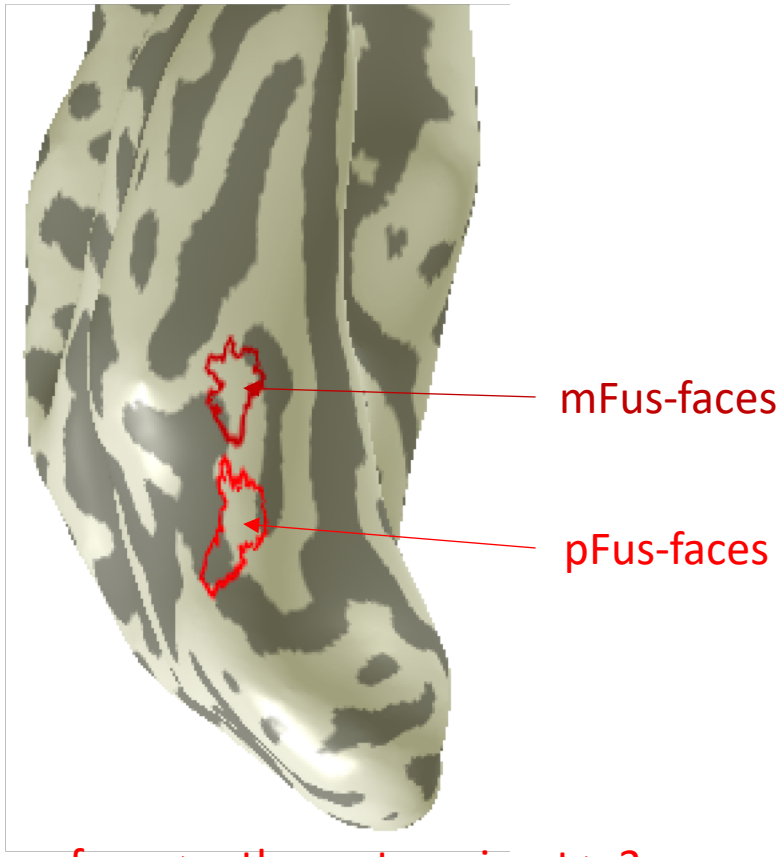


Grill-Spector et al. (2017), Annual Review of Vision Science

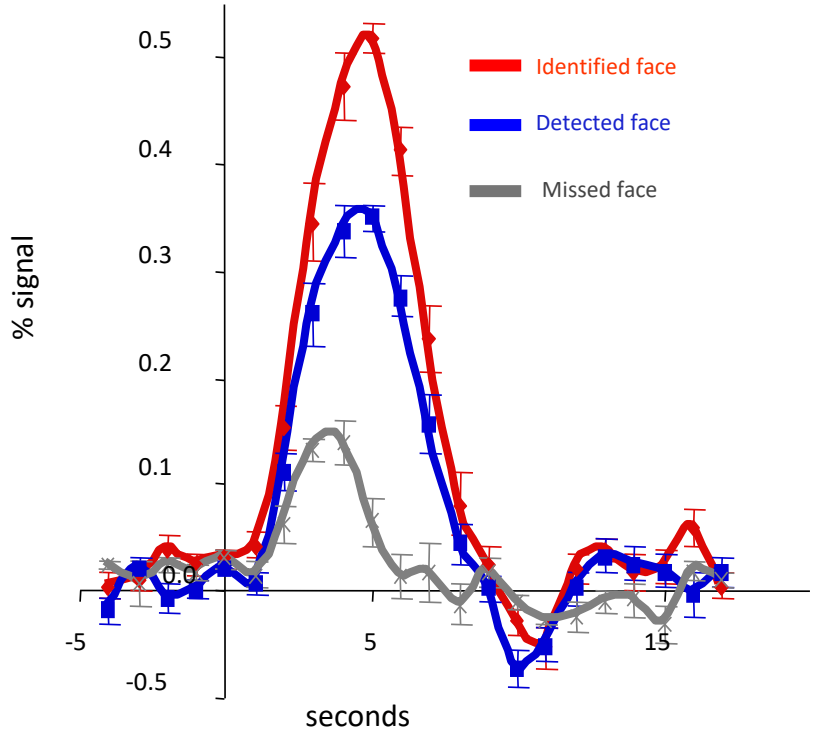


Davidenko et al. (2011), Human Brain Mapping

Responses in face-selective regions are correlated with the perception of faces

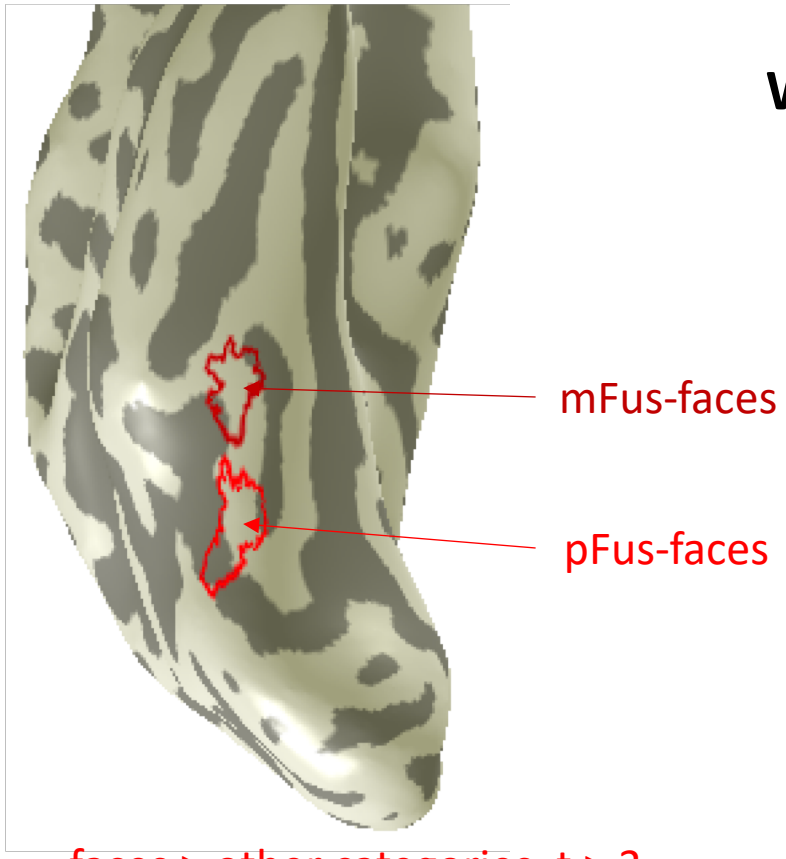


faces > other categories, t > 3



Grill-Spector et al. (2004), Nature Neuroscience

Studies so far largely investigated the role of category-selective regions in perceiving their preferred category



faces > other categories, t > 3

We were interested in building on these findings:

- **using carefully controlled stimuli**
- **that show ambiguous images**
- **to investigate between-category categorization**

Another domain-specific area is OTS-bodies, which shows the highest signal to bodies

faces > other categories, t > 3
limbs > other categories, t > 3

We were interested in building on these findings:

- **using carefully controlled stimuli**
- **that show ambiguous images**
- **to investigate between-category categorization**

Another domain-specific area is OTS-bodies, which shows the highest signal to bodies

We were interested in building on these findings:

- **using carefully controlled stimuli**
- **that show ambiguous images**
- **to investigate between-category categorization**
- **within face- and body-selective regions in VTC**

Research questions

(1) Does perceptual categorization correlate with neural responses in mFus-faces, pFus-faces and OTS-bodies?

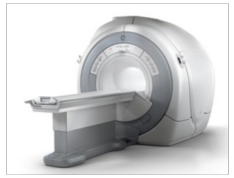
(2) Does the combination of multiple regions processing faces and hands explain perceptual categorization better than one by itself?

Morphed face silhouettes to hands silhouettes, chose 5 morphing steps

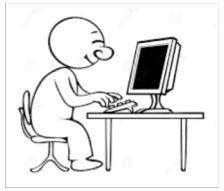
Experimental design

Stimulus calibration

fMRI blocked experiment



Behavioral categorization task



Model linking fMRI responses to behavioral categorization

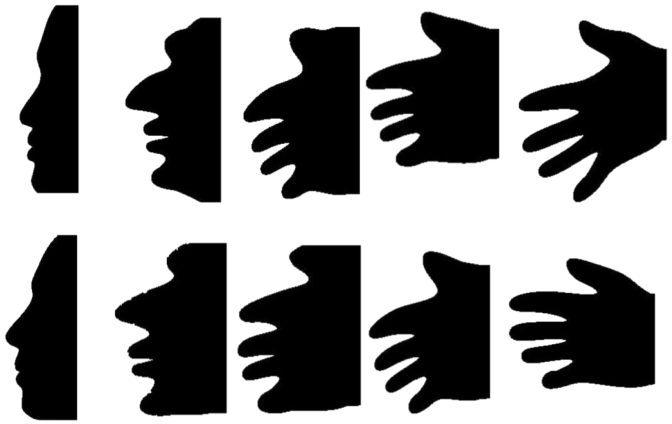
We calibrated the stimulus set on mechanical Turk in a large group of subjects

(1) Morph space normalization

N = 60 stimulus morphs

Purpose: perceptually equidistant morph steps

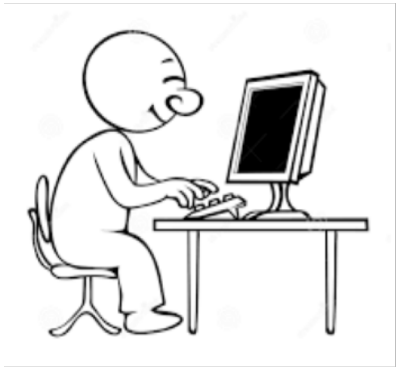
(2) Dissimilarity adjustment between stimulus exemplars

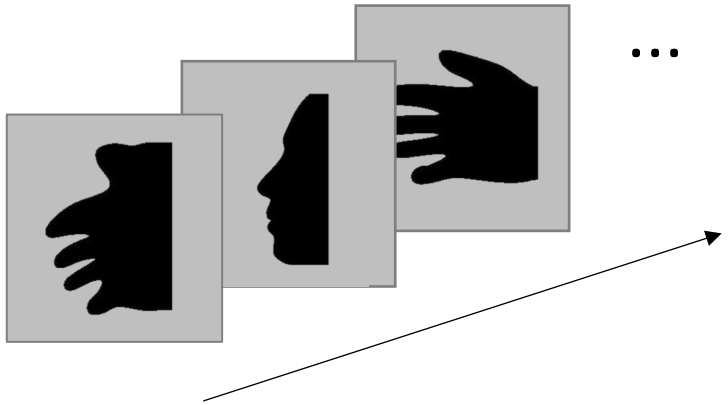


Purpose: constant dissimilarity within each morphing level

A new group of subjects (N = 14) performed a behavioral categorization task

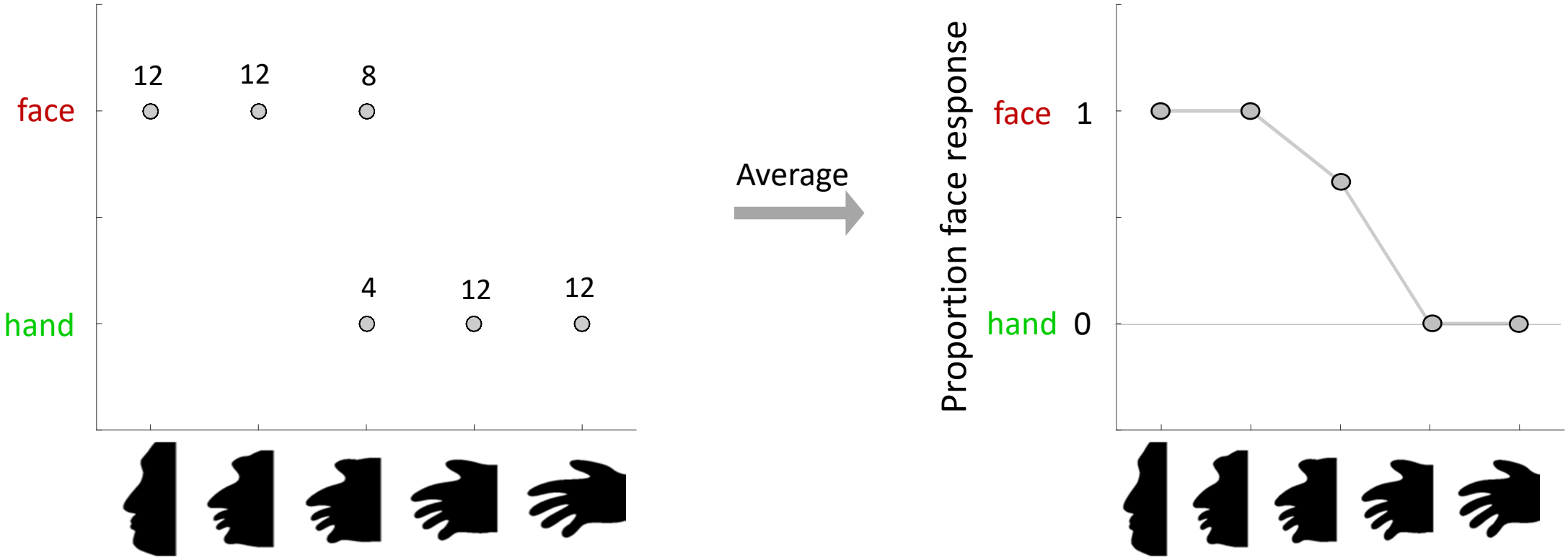
On each stimulus indicate: *Is this a face or a hand?*





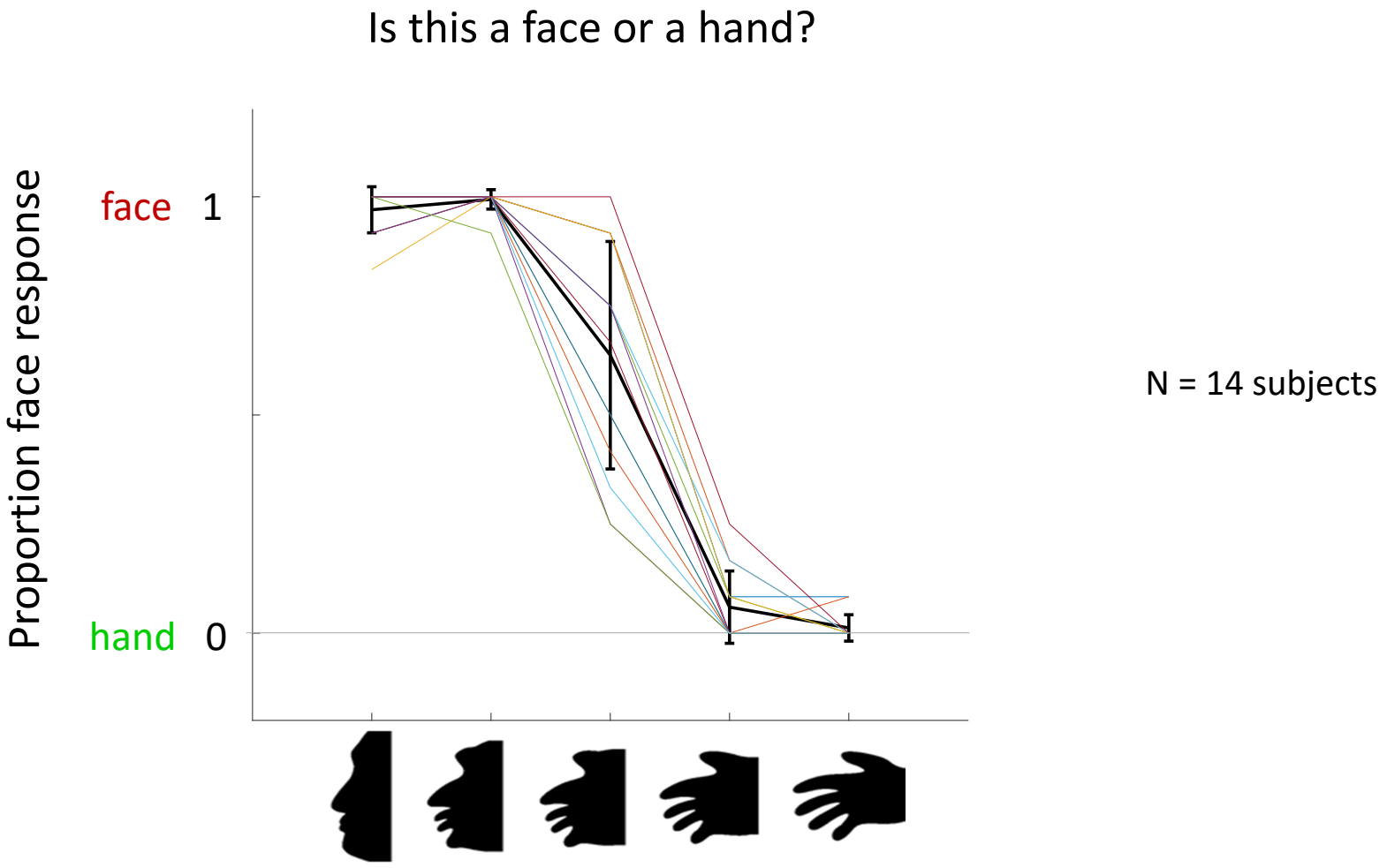
- 60 trials
- 12 different stimuli per morph level

How many times was each stimulus rated as a face or a hand? Example subject



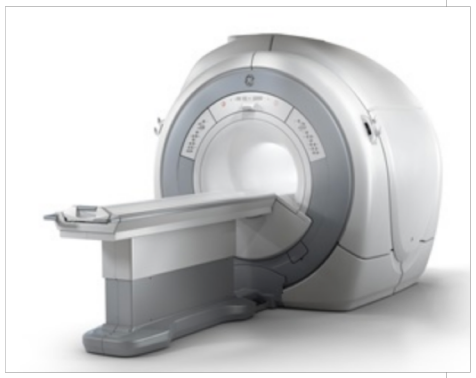
N = 60 trials

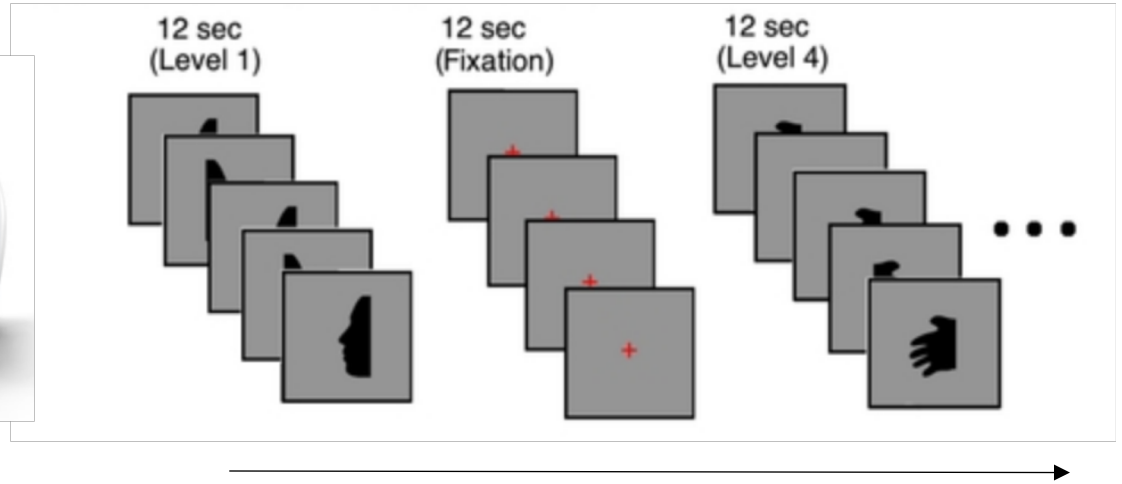
Middle morph level is perceived as ambiguous, category boundary varies across subjects



N = 14 subjects

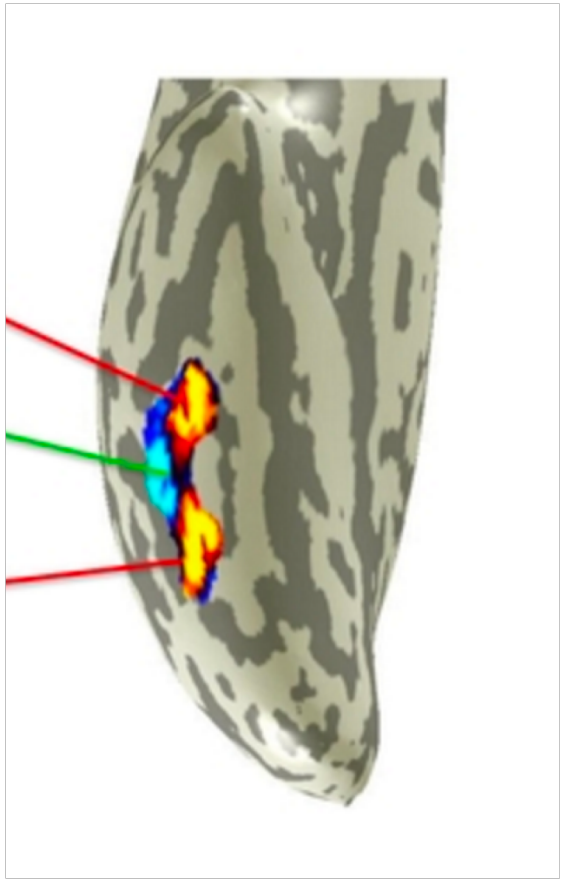
The same subjects (N = 14) participated in an fMRI blocked experiment



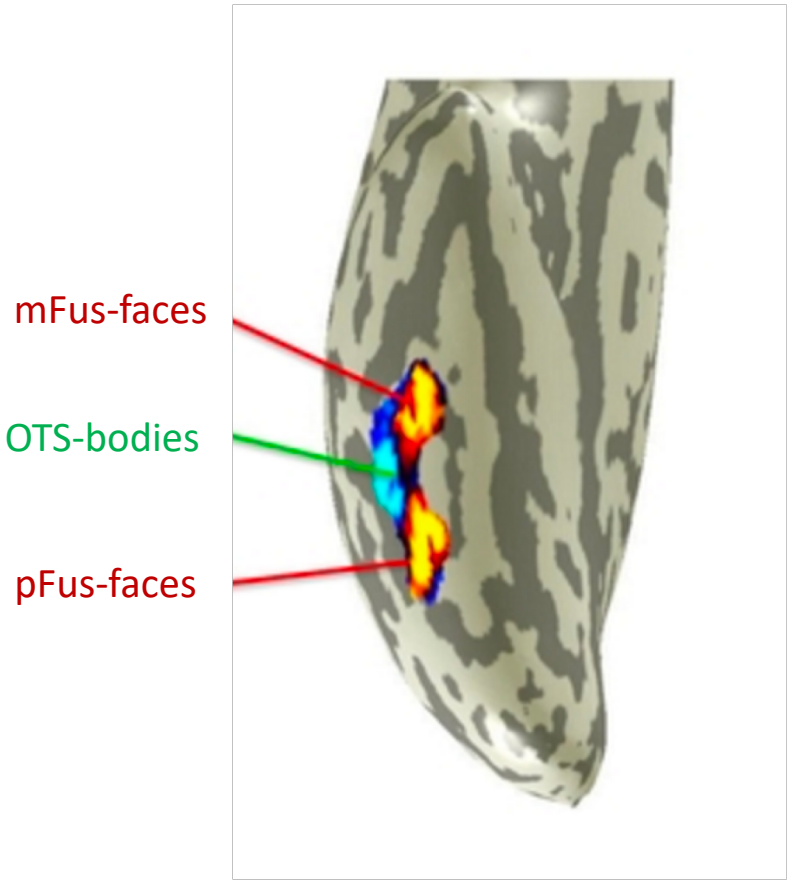


- 1-back task with fixation, indicating when two images are the same
- morph levels were pseudorandomized
- 4 runs, 126 time points each
- 2.4 mm isovoxel

We first localized face- and body-selective regions in VTC in each subject using an independent localizer



We first localized face- and body-selective regions in VTC in each subject using an independent localizer, and then extracted responses to the morphed stimuli from each region

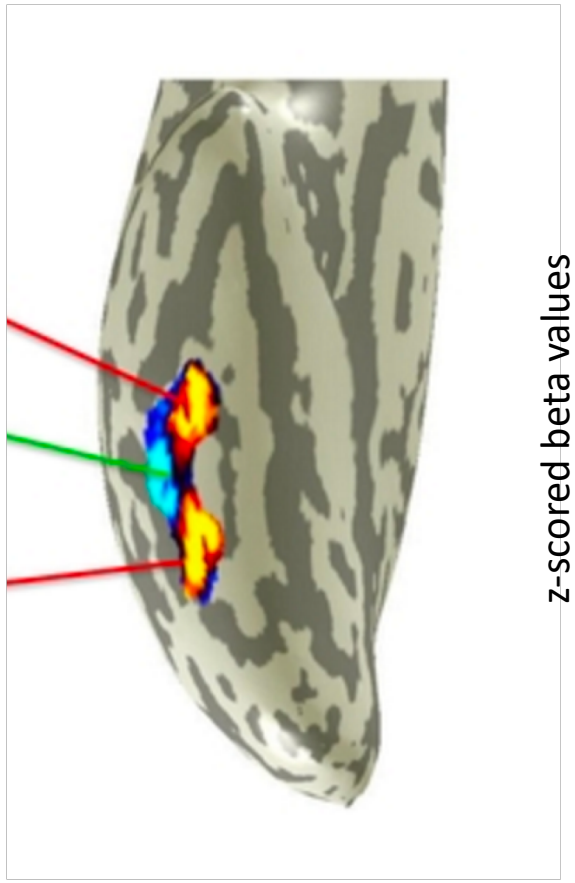


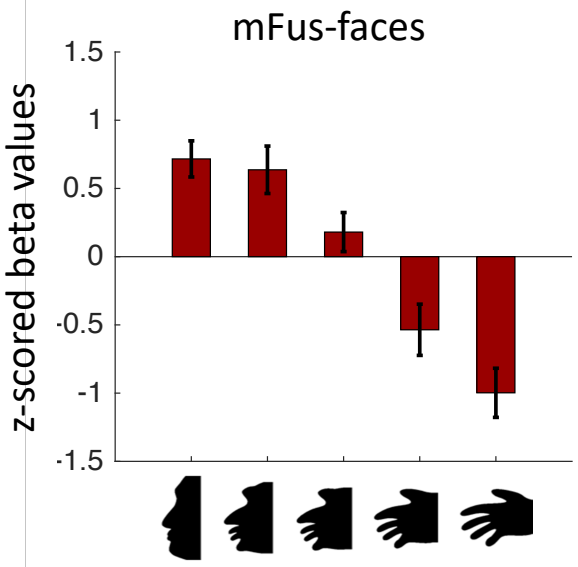
mFus-faces responds similarly to morphs 1,2 which are categorized as faces, then responses gradually decrease

mFus-faces

OTS-bodies

pFus-faces



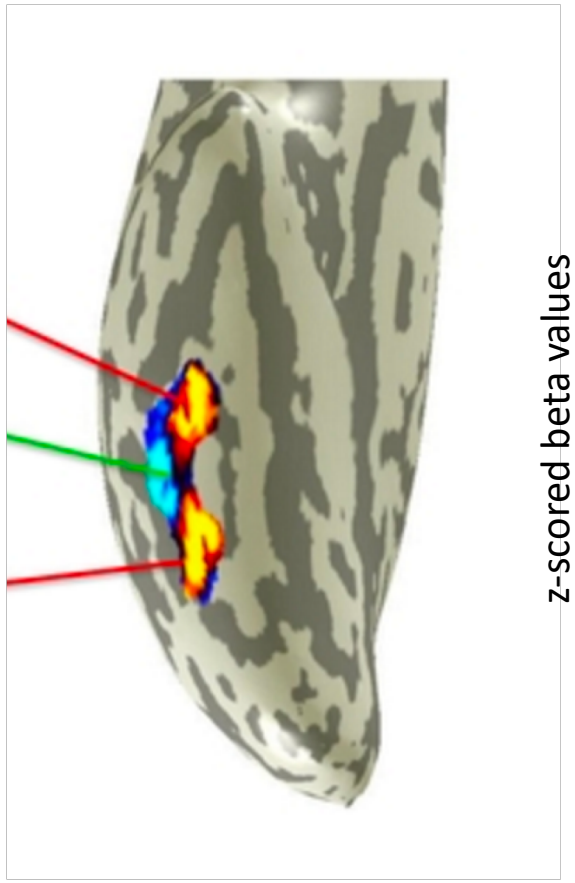


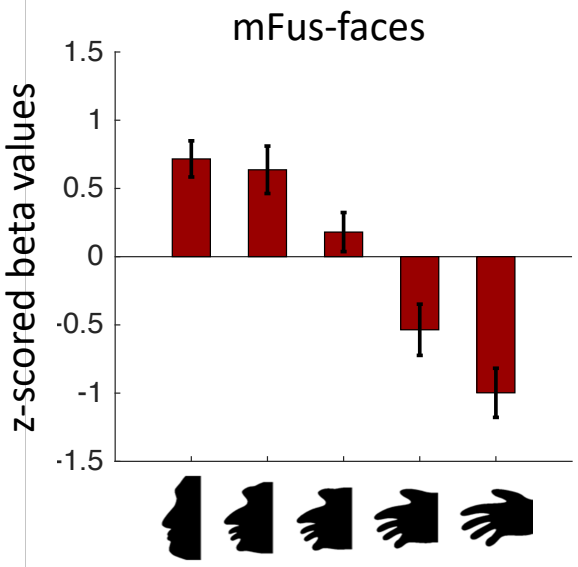
pFus shows a similar pattern

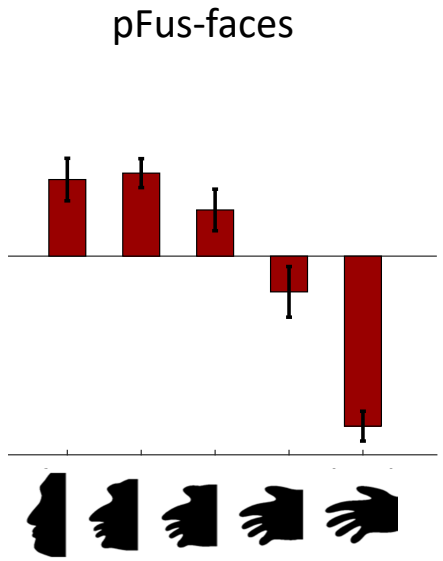
mFus-faces

OTS-bodies

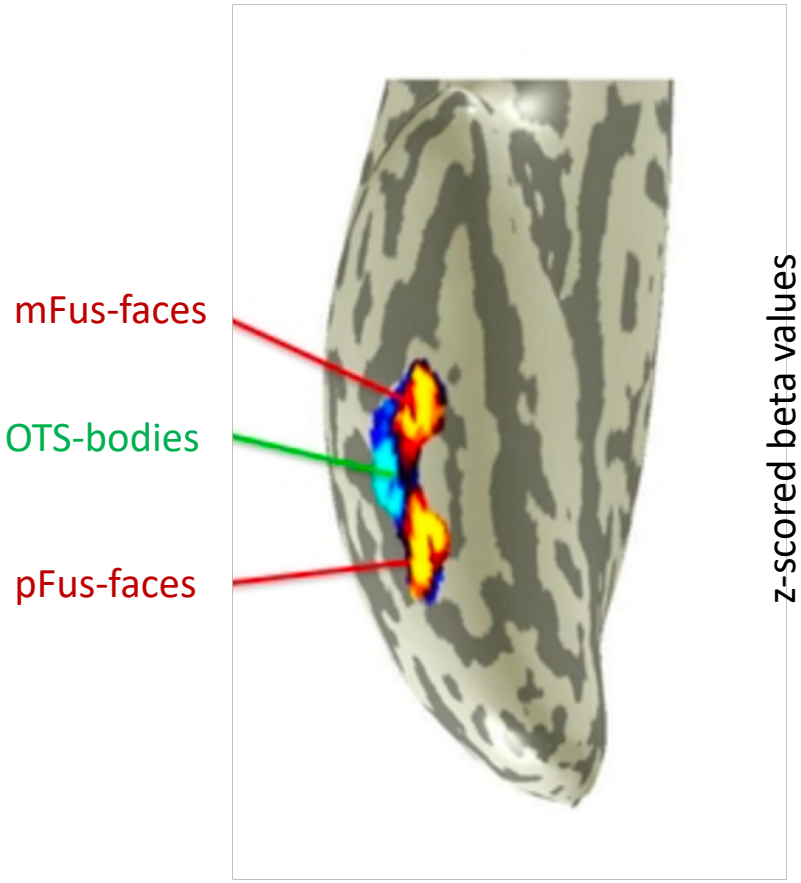
pFus-faces

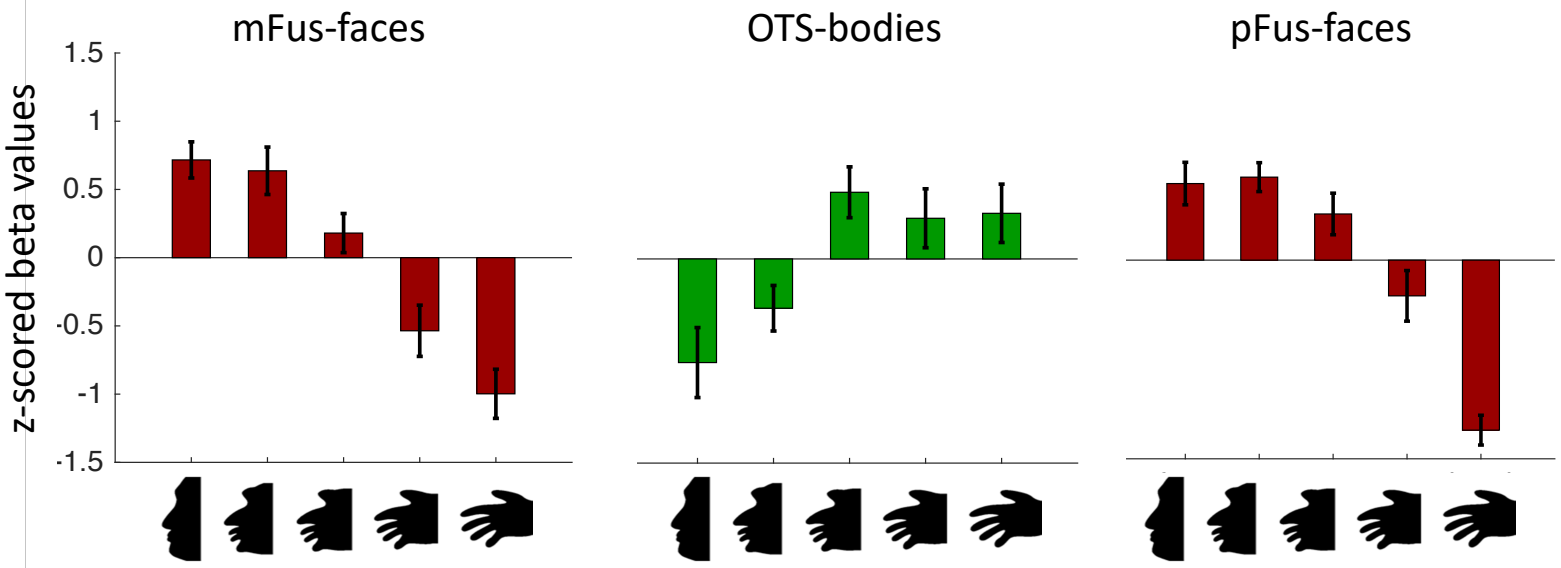






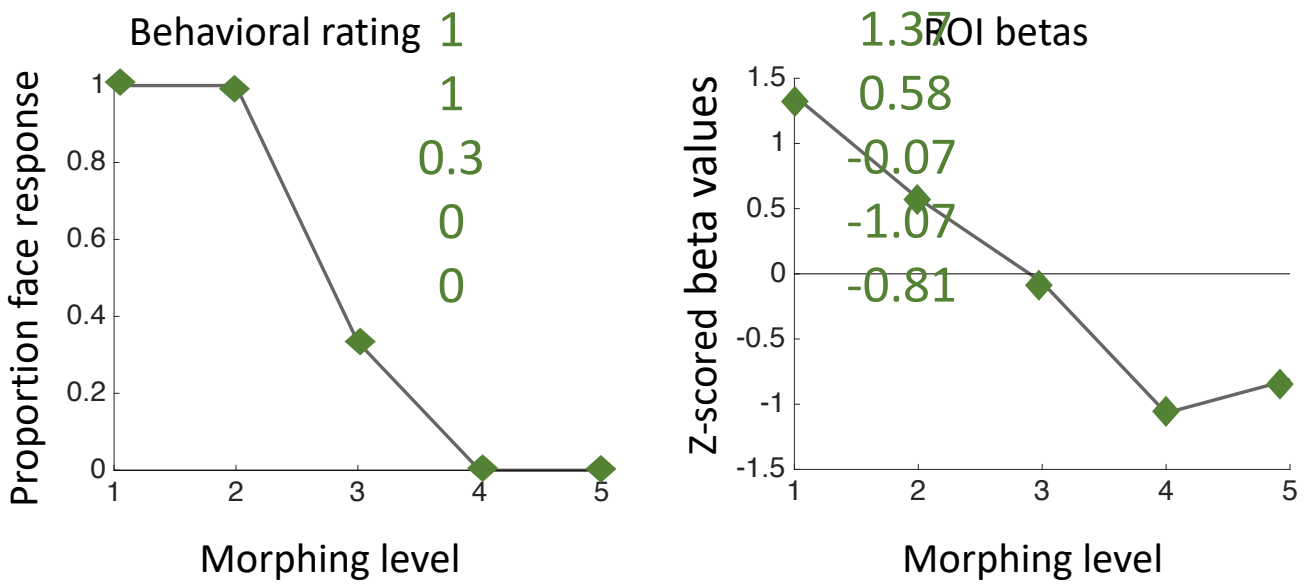
In contrast, OTS-bodies responds similarly to morphs 3-5, and gradually less to morphs 1,2



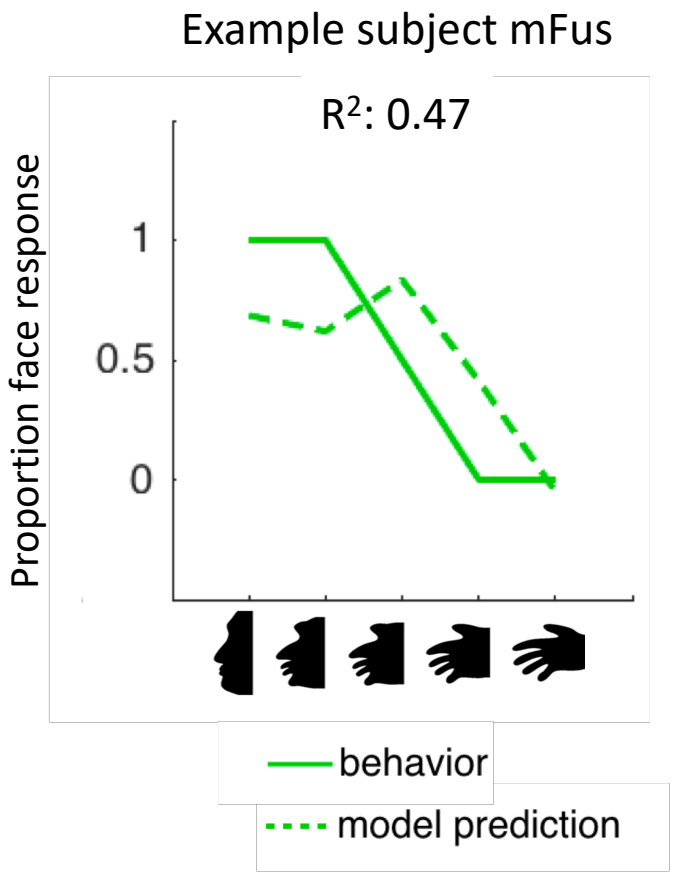


Does perceptual categorization correlate with neural responses in mFus-faces, pFus-faces and OTS-bodies?

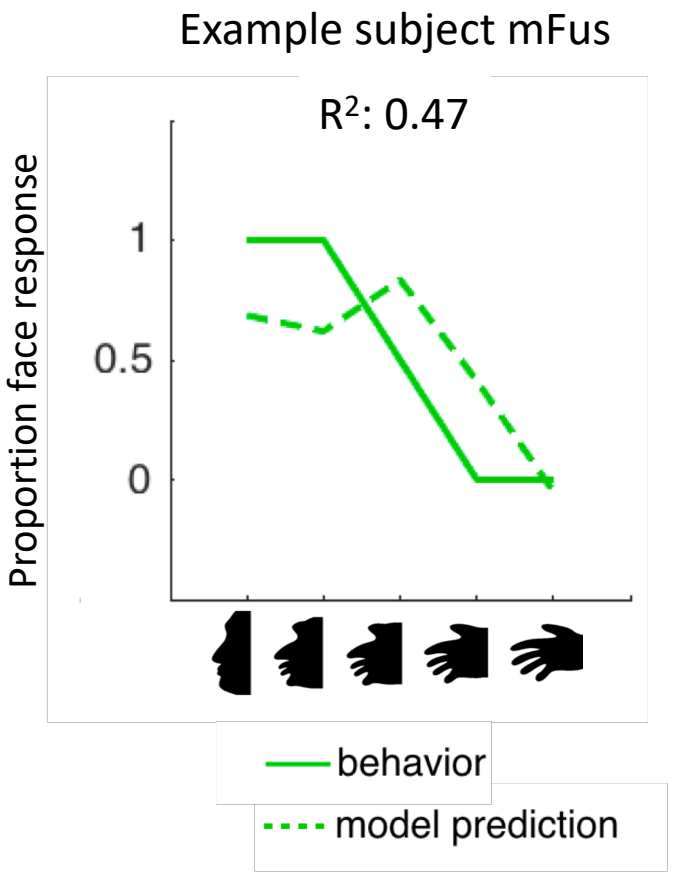
$$y = \alpha + \beta x$$

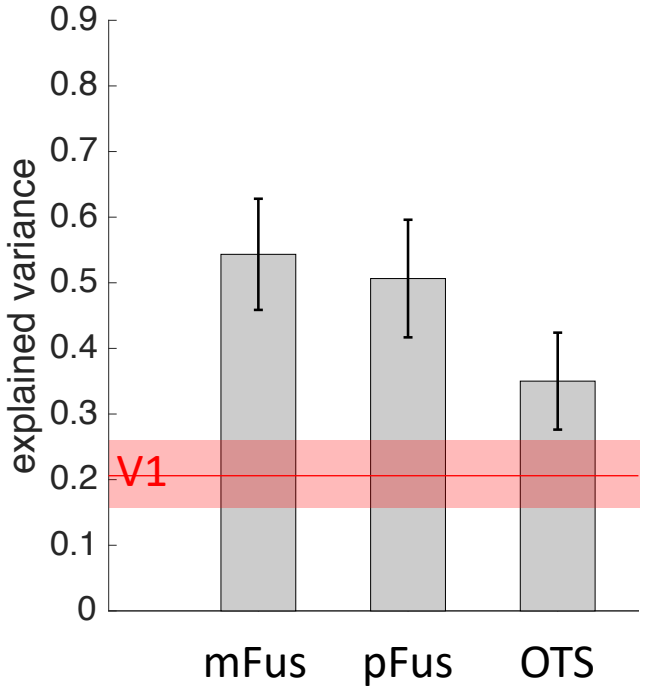


Example subject's mFus-faces data in relation to their behavioral categorization task

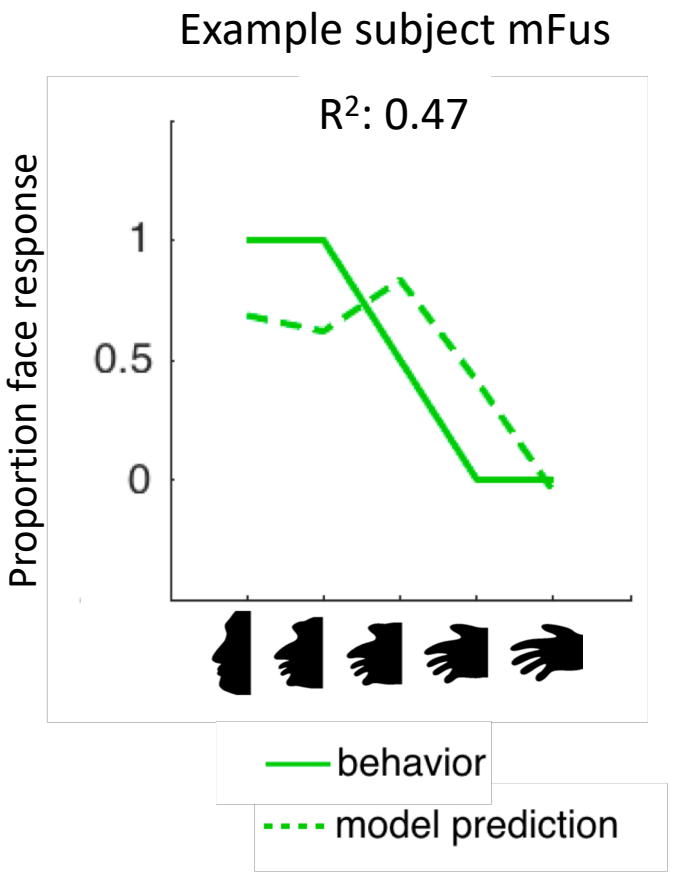


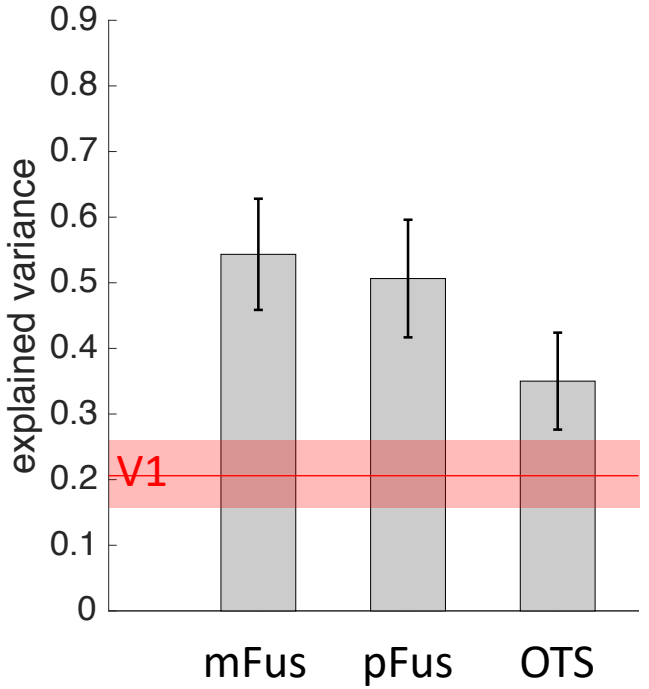
mFus-faces, pFus-faces and OTS-bodies are significantly higher correlated to behavioral categorization than control area V1





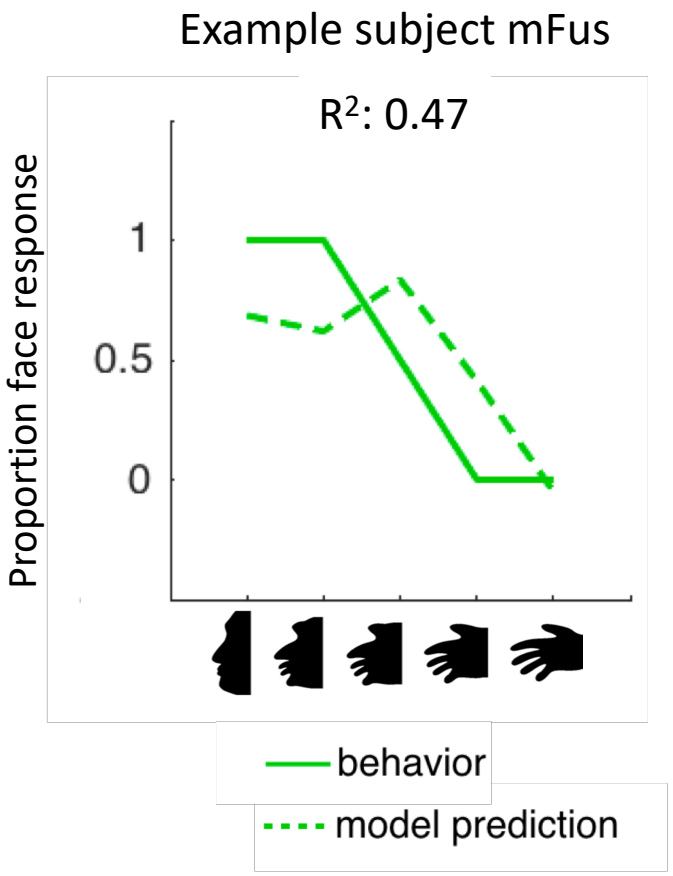
mFus-faces, pFus-faces and OTS-bodies are significantly higher correlated to behavioral categorization than control area V1

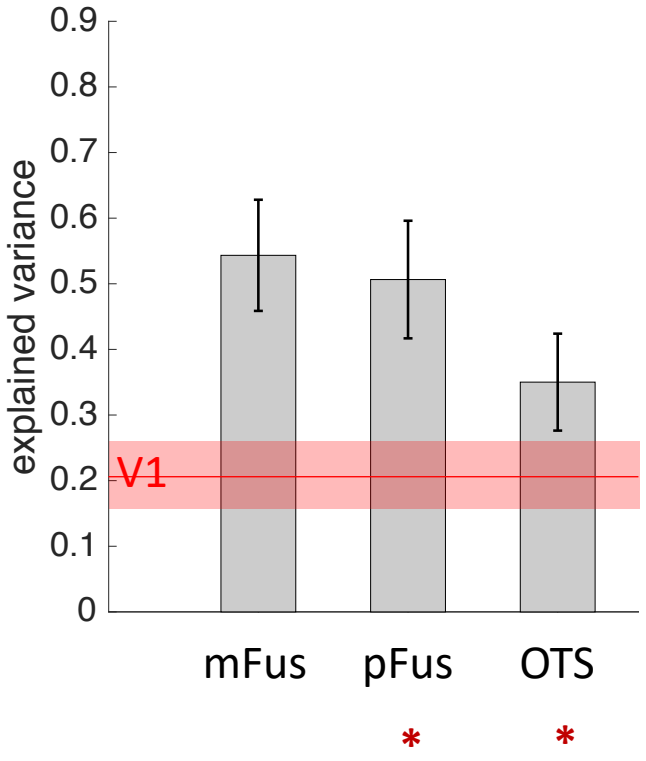




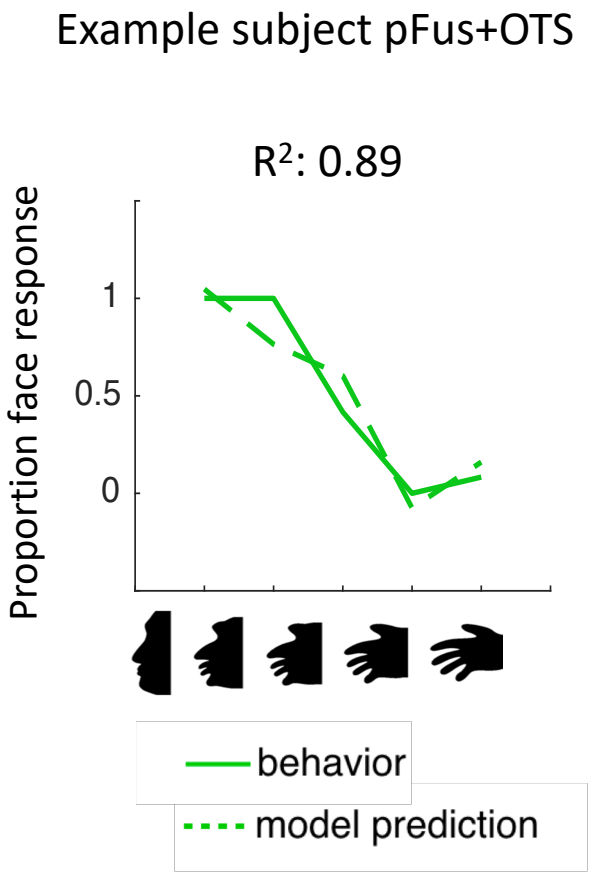
Stepwise linear regression

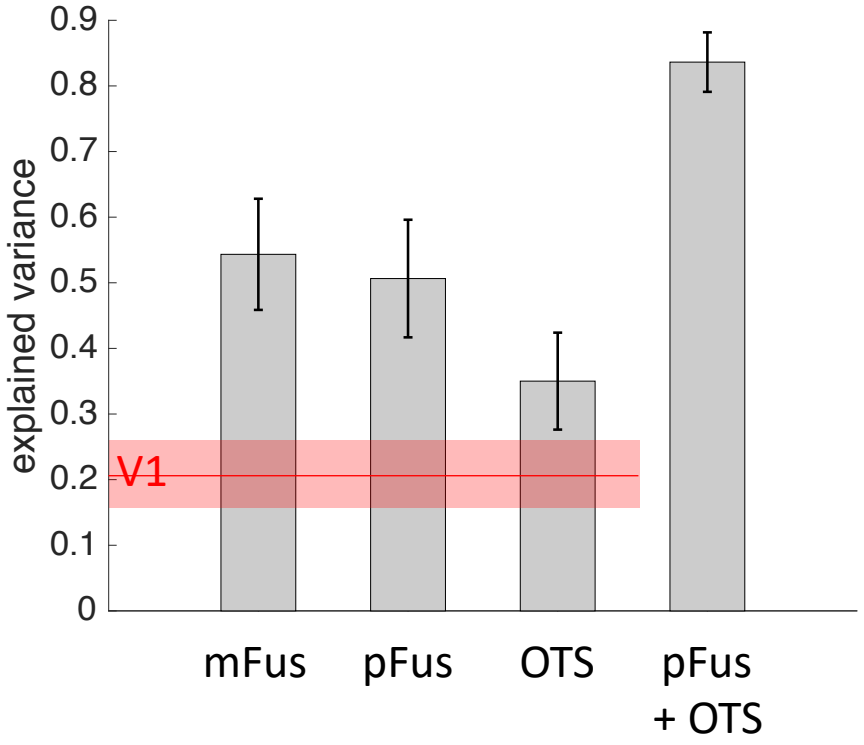
Stepwise linear regression indicates that both pFus and OTS are significant predictors in explaining behavioral categorization



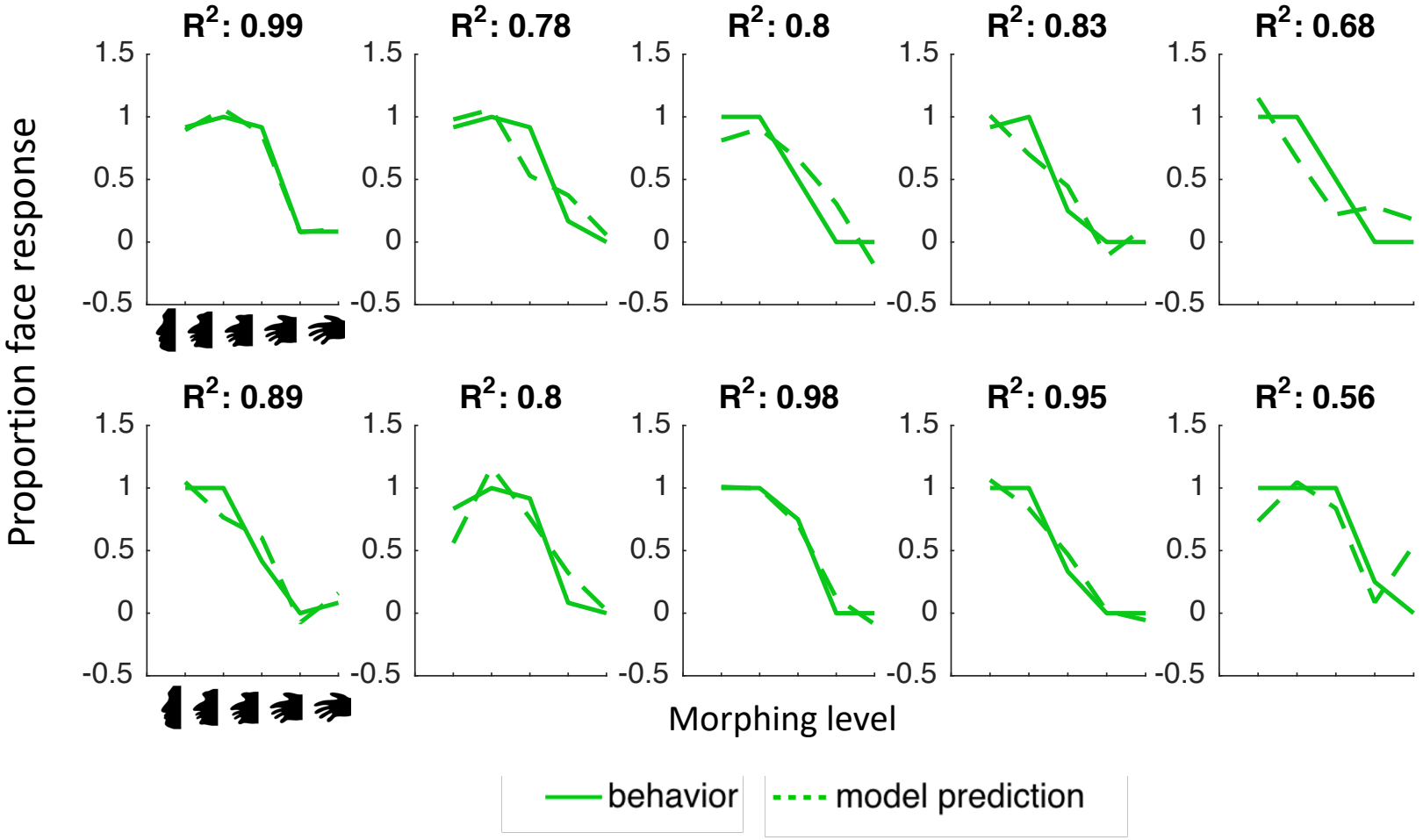


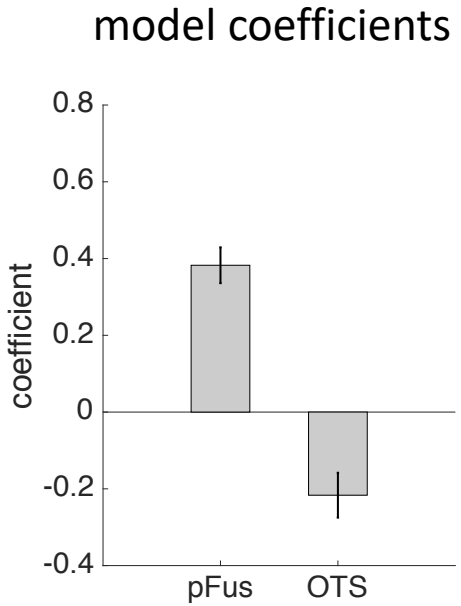
A linear model using pFus and OTS as predictors explains on average 82% of the variance of behavioral categorization



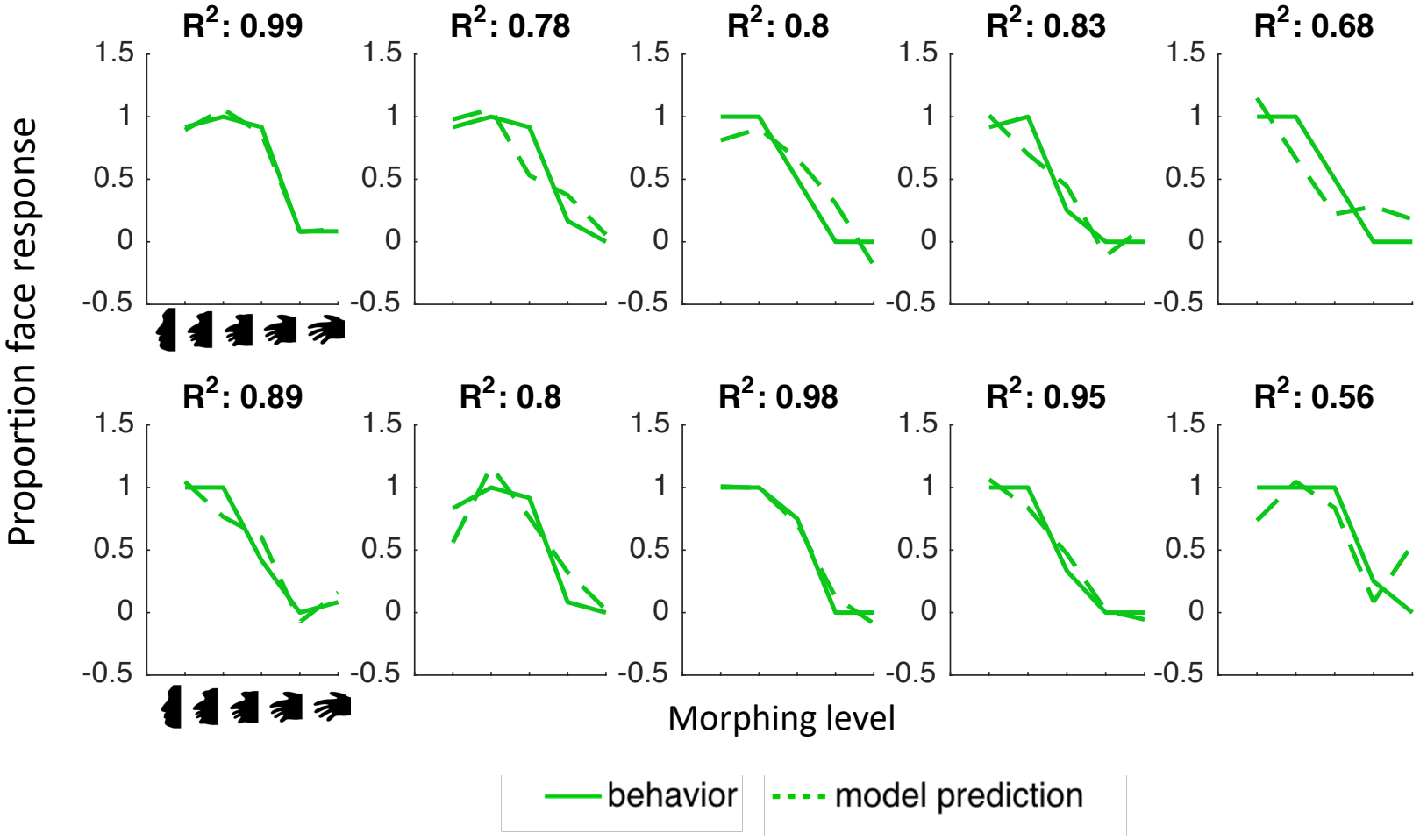


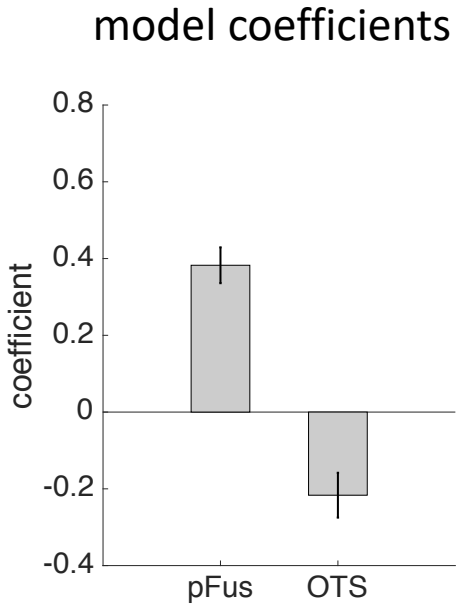
A linear model using pFus and OTS as predictors explains on average 82% of the variance of behavioral categorization



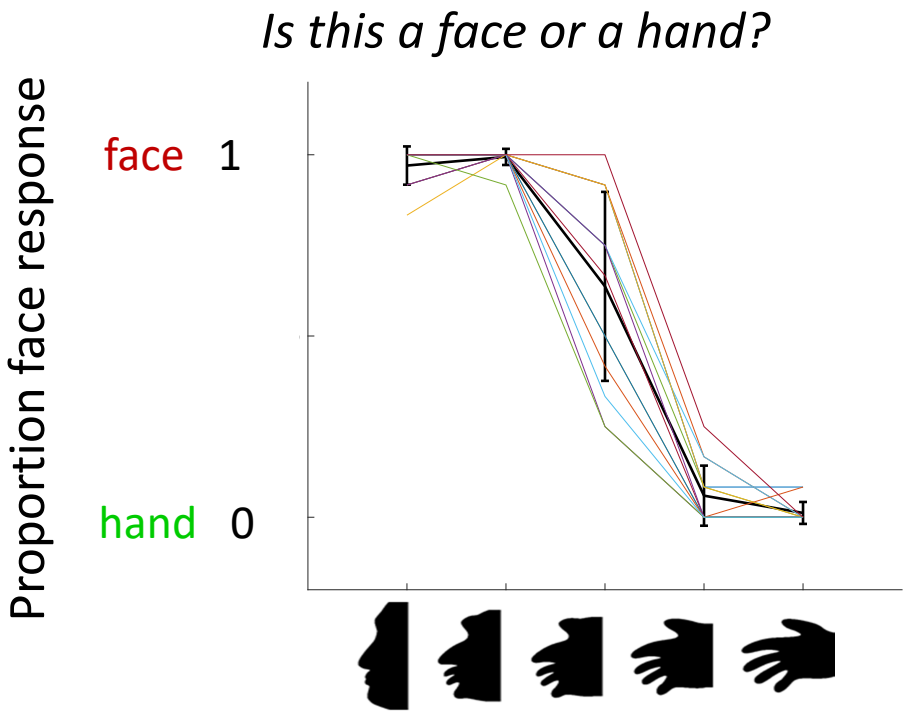


We can best explain categorization behavior in each subject by taking the relative contribution of pFus-faces and OTS-bodies.





Can we build a general model?



Can we build a general model to predict your behavioral categorization response?

Training model on N-1 subjects

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2$$

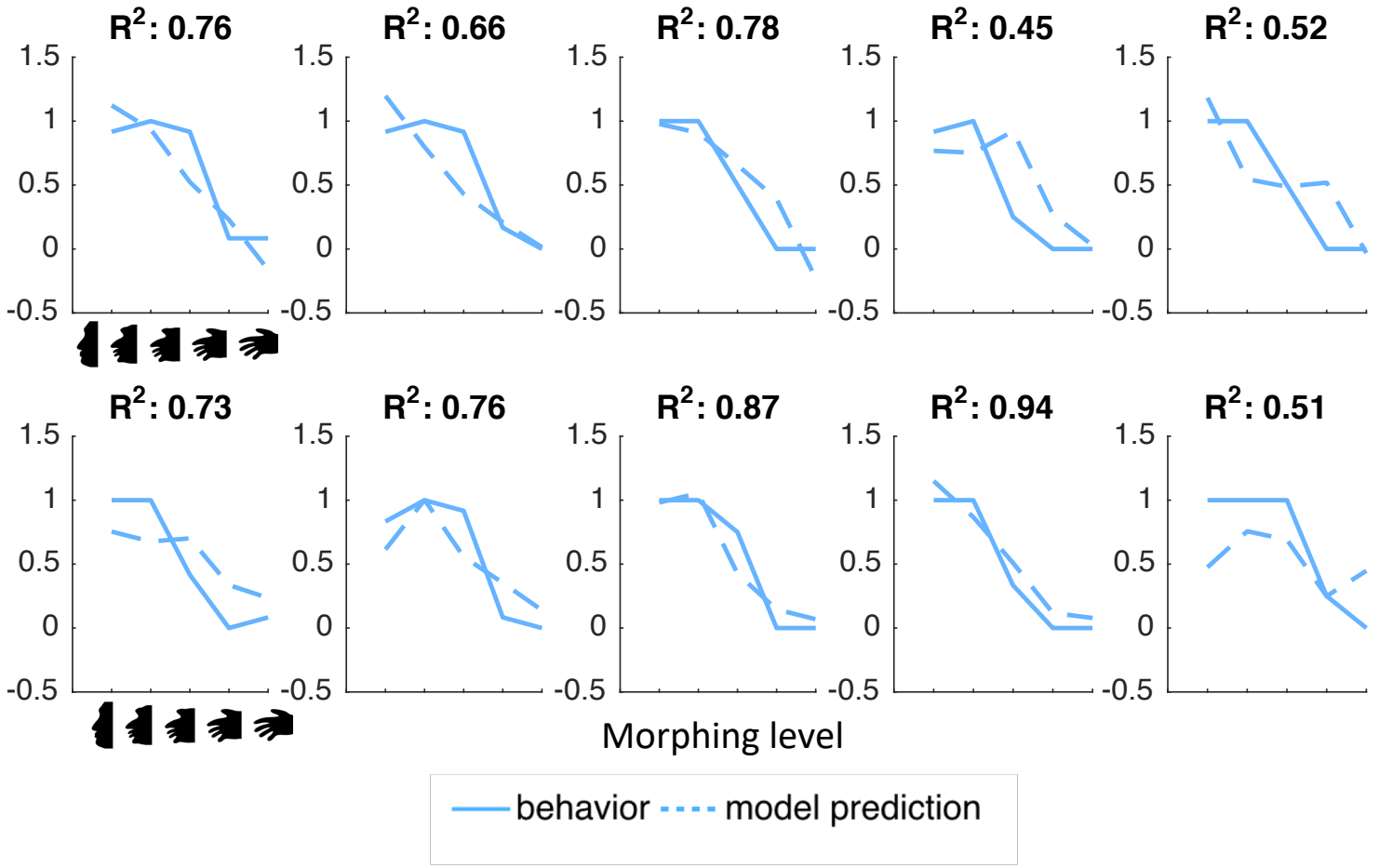
subject 1 **behavior**	subject 1 **pFus** betas	subject 1 **OTS** betas
subject 2 **behavior**	subject 2 **pFus** betas	subject 2 **OTS** betas
...	...	...
subject N-1 **behavior**	subject N-1 **pFus** betas	subject N-1 **OTS** betas

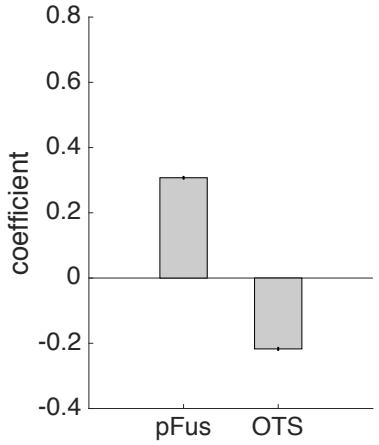
α, β_1, β_2

Testing on Nth subject

subject N behavior = α + β_1 . subject N **pFus** betas + β_2 subject N **OTS** betas

We can build a model that predicts perceptual categorization in new subjects from their brain responses;
Model explains 70% of the variance in new subjects





Conclusion

(1) Perceptual categorization of ambiguous face-hand stimuli is better predicted by differential responses across a face and a body region than responses of each of these regions separately.

(2) These differential responses explain ~82% variance of perceptual categorization in individual subjects

(3) Elucidating the relationship between brain activity and perceptual categorization in individual subjects enabled us to build a general model that predicts ~70% of the variance of perceptual categorization in new subjects, solely based on their neural data

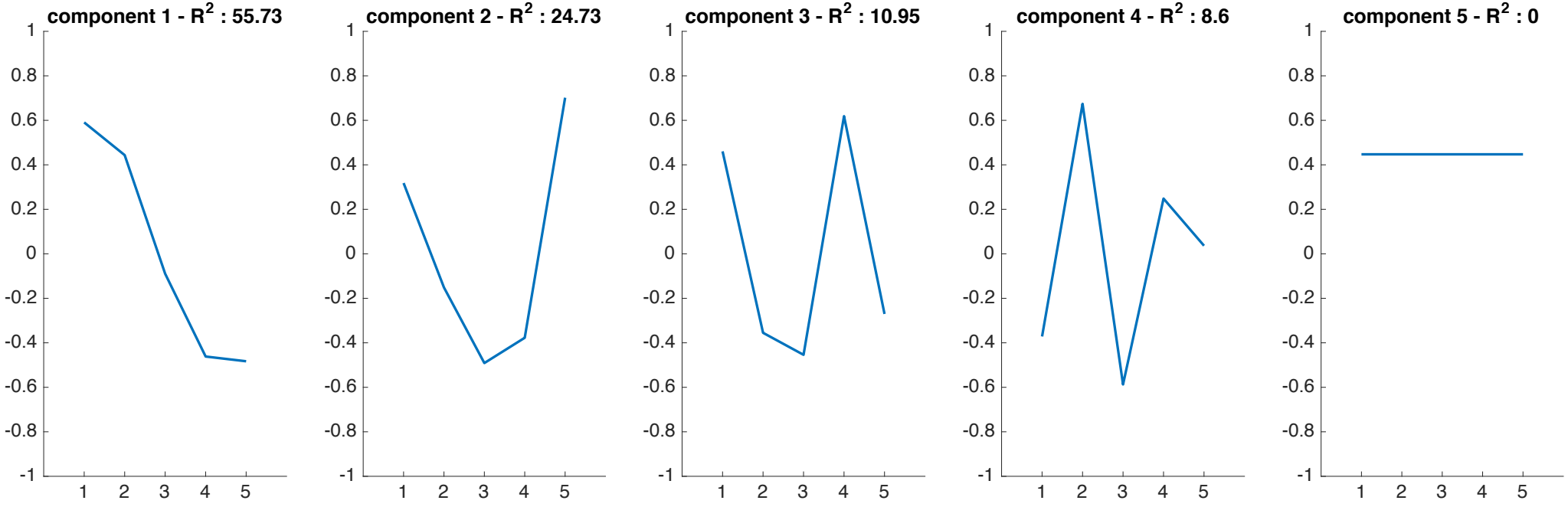
Thank you!

Kalanit Grill-Spector

Kevin S. Weiner

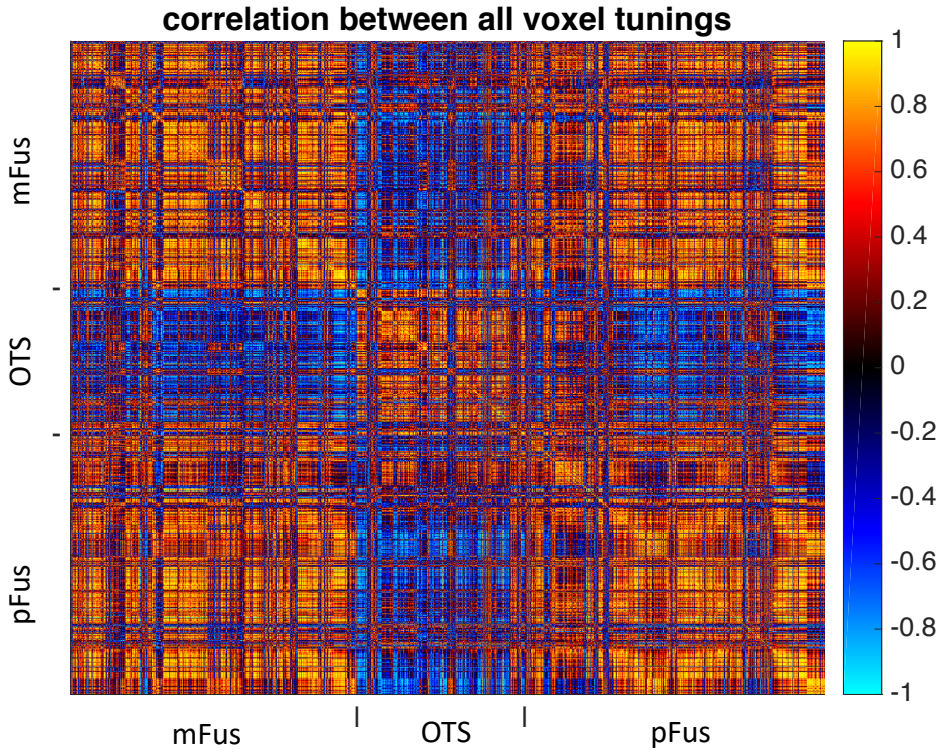
Nicolas Davidenko

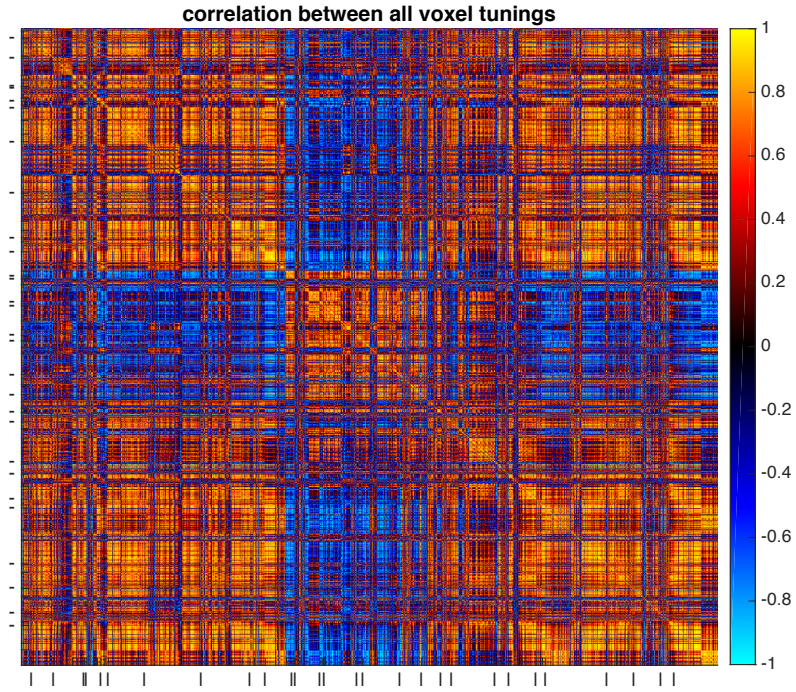
PCA results



81 % variance explained by first two components

Correlation across voxel tunings of mFus, OTS, pFus





Same matrix as on the left, but subjects are marked on axes

=> There is variability in tuning functions within an ROI and variability across subjects, but the dominant difference is between face and body selective-region's tuning curves