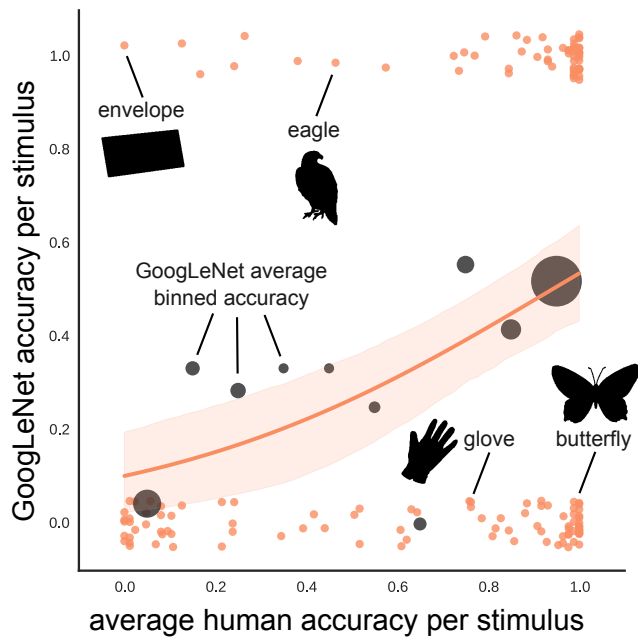# Exp. 1: Recognition from silhouettes

Humans and GoogLeNet perform similarly on object recognition using 147 silhouettes of common objects

5 shallow models (no training)
3 HMAX models (trained on natural scenes)
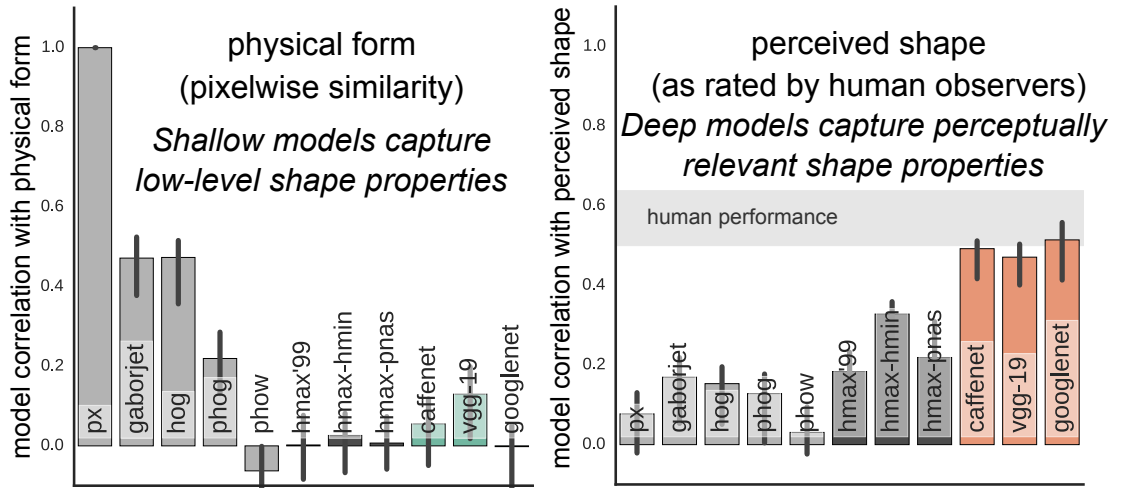3 deep convolutional networks (trained on the ImageNet for object recognition)

Exp. 2: Non-accidental properties

Like humans, deep nets find non-accidental changes in shape more different than metric changes
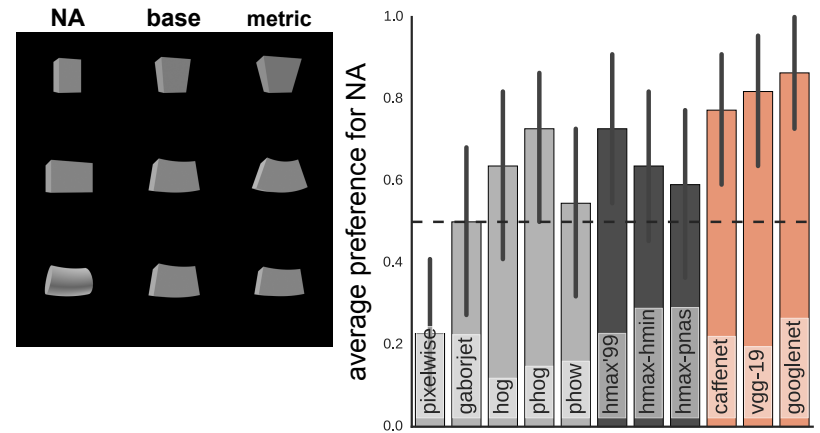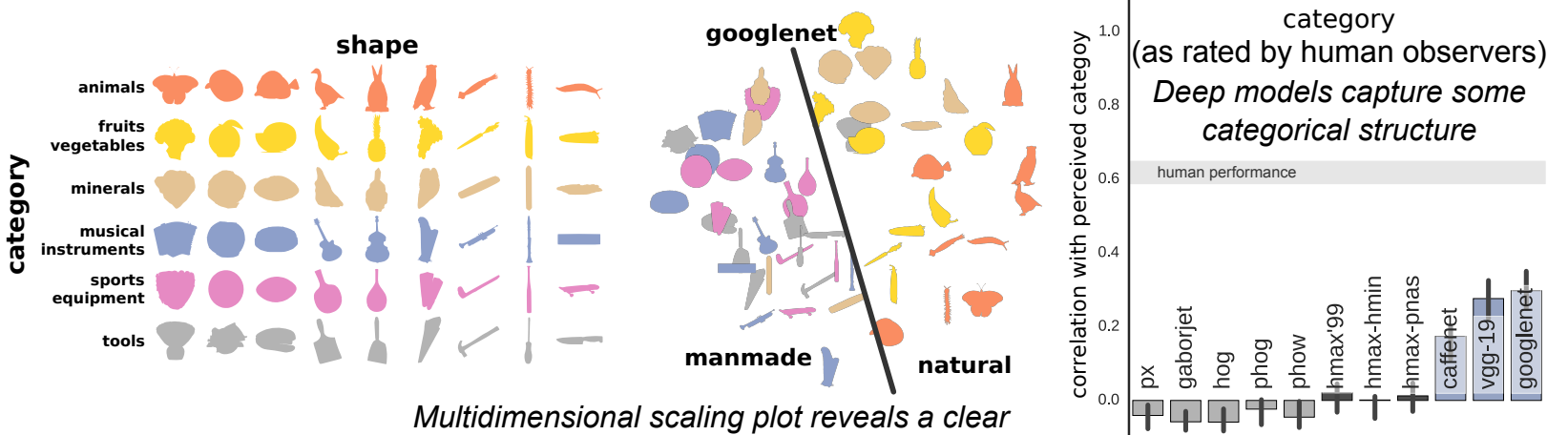

Exp. 3: Physical vs. perceived shape

Shapes can be compared in terms of their pixelwise similarity or their perceived similarity as judged by humans. We compared the resulting Representational Similarity Matrices between humans and models.

stimuli

physical form
(pixelwise similarity)

Shallow models capture low-level shape properties

perceived shape
(as rated by human observers)

Deep models capture perceptually relevant shape properties

Exp. 4: The role of category when dissociated from shape

shape

category

animals
fruits vegetables
minerals
musical instruments
sports equipment
tools

googlenet

manmade   **natural**

category
(as rated by human observers)

Deep models capture some categorical structure

human performance

Multidimensional scaling plot reveals a clear separation between manmade and natural objects in deep nets. Note that the models were never trained on such superordinate categories.